Cell, Vol. 63, 167–174, October 5, 1990, Copyright © 1990 by Cell Press

Identification of a Ligand for the c-*kit* Proto-Oncogene

Douglas E. Williams,* June Eisenman,* Allison Baird,* Charles Rauch,† Kirk Van Ness,† Carl J. March,† Linda S. Park,‡ Unja Martin,‡ Diane Y. Mochizuki,* H. Scott Boswell,§ Gem S. Burgess,§ David Cosman,‖ and Stewart D. Lyman‖

* Department of Experimental Hematology
† Department of Protein Chemistry
‡ Department of Biochemistry
‖ Department of Molecular Biology
Immunex Corporation
51 University Street
Seattle, Washington 98101
§ Department of Medicine (Hematology/Oncology)
Indiana University School of Medicine
Veterans Administration Hospital
Indianapolis, Indiana 46223

Summary

We report the purification and N-terminal amino acid sequence of a novel mast cell growth factor, termed MGF, from the supernatants of a murine stromal cell line. A panel of interleukin 3–dependent cell lines were screened for responsiveness to partially purified MGF in [^{3}H]thymidine incorporation assays; proliferative stimulation of these cells in response to MGF correlated with expression of mRNA for the c-*kit* proto-oncogene. MGF was shown to be a ligand for c-*kit* by cross-linking ^{125}I-labeled MGF to c-*kit*-expressing cells with subsequent immunoprecipitation of the complex with antiserum specific for the C-terminus of c-*kit*. This establishes MGF as a ligand for the c-*kit* protein.

Introduction

Mice with mutations of the steel (*Sl*) or dominant spotting (*W*) locus display a similar phenotype, characterized by a reduction of pluripotent hematopoietic stem cells (PHSCs), anemia, a deficiency of mast cells, and defects in gametogenesis and pigmentation (Russell, 1979; Mintz and Russell, 1957; Mayer and Green, 1968). The hematopoietic defect in *W* mutant mice is intrinsic to the PHSC (Russell, 1979; Russell et al., 1956; Russell and Bernstein, 1968; Murphy et al., 1973), whereas that of the *Sl* mouse is the result of a microenvironmental aberration (Russell, 1979; Bernstein, 1970; Harrison and Russell, 1972; Fried et al., 1973). Recent studies have shown that the *W* locus on murine chromosome 5 encodes the c-*kit* proto-oncogene (Chabot et al., 1988; Geissler et al., 1988). This proto-oncogene encodes a tyrosine kinase receptor (Yarden et al., 1987) that is most closely related structurally to c-*fms* (Qiu et al., 1988), the receptor for macrophage colony-stimulating factor (CSF-1) (Sherr et al., 1985), and to the PDGF receptor (Yarden et al., 1986). Previous studies have shown that bone marrow from *Sl* mice transplanted into *W* recipients resulted in a hematologically normal phenotype (Bernstein et al., 1968) and has led to the hypothesis that the gene product of the *Sl* locus is a ligand for c-*kit* (Chabot et al., 1988; Geissler et al., 1988).

Long-term growth of primitive hematopoietic cells in vitro can be done using the Dexter-type culture system (Dexter et al., 1971). Both establishment and maintenance of these cultures require a complex stromal cell layer that provides the microenvironmental signals necessary for hematopoiesis. The colony-stimulating factors (CSFs) granulocyte-macrophage CSF, granulocyte CSF, macrophage CSF, and multi-CSF (interleukin 3; IL-3) do not appear to be important mediators in these cultures (Gualteri et al., 1987; Eliason et al., 1988; Li and Johnson, 1985). It is therefore likely that some other stromal cell–derived factor is responsible for regulating cell production in these cultures. This idea is supported by the fact that Dexter cultures initiated with bone marrow from *Sl/Sl^d* or *W/W^v* bone marrow gave rise to defective cultures, consistent with the in vivo phenotype, and the addition of *Sl/Sl^d* bone marrow to preexisting *W/W^v* stromal layers resulted in long-term hematopoiesis (Dexter and Moore, 1977; Boswell et al., 1987).

We have previously described a pair of unique stromal cell clones derived from Dexter cultures of *Sl/Sl^d* bone marrow and their normal littermates (Boswell et al., 1990). The cell line from the mutant mice was unable to maintain the viability of a panel of IL-3-dependent mast cell lines when cocultured, whereas the cell line derived from normal mice supported the survival of two mast cell lines in the absence of exogenous growth factors (Boswell et al., 1990). A growth factor for the mast cell lines was detected in concentrated supernatants from the normal stromal cell line, and this activity was not attributable to IL-1 through IL-7, CSFs, leukemia inhibitory factor (Boswell et al., 1990), or IL-9 (D. E. W., unpublished data). No mast cell stimulatory activity could be detected in supernatants from the *Sl/Sl^d* cell line. These observations were significant because one of the primary defects in *Sl* mice is in the mast cell lineage (Kitamura and Go, 1978; Kitamura et al., 1978).

Several mast cell lines were used to assay this bioactivity produced by the normal stromal cell line. We now report the purification and N-terminal amino acid sequence analysis of a novel stromal cell–derived murine hematopoietin that stimulates the proliferation of a select group of IL-3-dependent mast cell lines. Proliferative responses to this growth factor correlated with c-*kit* mRNA expression, and the purified, radiolabeled protein could be cross-linked to responder cells and subsequently immunoprecipitated with antiserum that recognizes the C-terminus of the c-*kit* gene product. This establishes this new growth factor, termed mast cell growth factor (MGF), as a ligand for the c-*kit*-encoded protein.

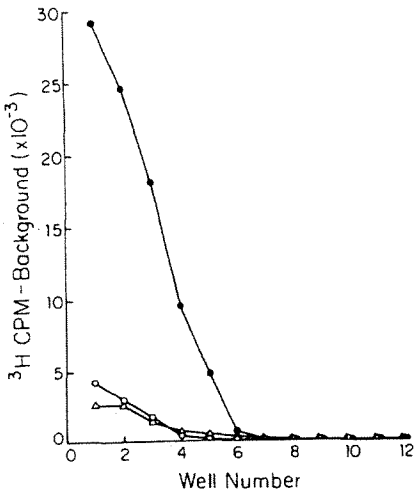

Figure 1. Stimulation of [^{3}H]Thymidine Incorporation by Partially Purified +/+ Conditioned Medium with Various Cell Lines

Shown are results with the MC-6 (filled circles), H7 (open circles), and FDC-P2-1D (triangles) cell lines. Data depict one representative titration and are shown as ^{3}H cpm minus background (y axis) versus sample dilution (x axis). Partially purified +/+ conditioned medium was present at a concentration of 25% (v/v) in the first well and serially diluted in 2-fold increments. Not shown are the results with the DA-1, FDC-P2, and 32D cell lines, which did not show significant [^{3}H]thymidine incorporation above that observed with medium alone.

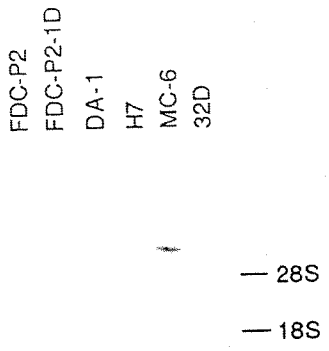

Figure 2. Northern Blot Analysis of Polyadenylated RNA from the FDC-P2, FDC-P2-1D, DA-1, H7, MC-6, and 32D Cell Lines

The blot shows the expression of the 5.5 kb c-*kit* mRNA transcript in MC-6, H7, and FDC-P2-1D cells. Blots were hybridized with a ^{32}P-labeled oligonucleotide probe as described in Experimental Procedures and subjected to autoradiography for 14 days. The positions of the 28S and 18S ribosomal RNAs are indicated.

Results

Correlation between c-*kit* mRNA Expression and Proliferation in Response to +/+ Conditioned Medium

We previously reported that medium conditioned by the +/+ cell line contains a hematopoietin (MGF) that stimulates the proliferation of two IL-3-dependent mast cell lines, H7 and MC-6 (Boswell et al., 1990). This proliferative activity could not be ascribed to IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, GM-CSF, CSF-1, G-CSF, LIF (Boswell et al., 1990), or IL-9 (D. E. W., unpublished data). Because a similar activity was not detected in supernatants from a stromal cell line derived from a steel mouse (Boswell et al., 1990) and because of the mast cell phenotype of the responsive cells, we postulated that this growth factor might be a ligand for c-*kit*.

We initially tested this hypothesis by examining a panel of IL-3-dependent cell lines for their capacity to proliferate in response to partially purified (S-Sepharose and DEAE) MGF. H7, MC-6, and FDC-P2-1D demonstrated enhanced proliferation in the presence of MGF, whereas FDC-P2, 32D, and DA-1 cells did not (Figure 1). These same cell lines were also examined for expression of c-*kit* mRNA by Northern blot analysis. Only the cell lines that proliferated in response to MGF (H7, MC-6, FDC-P2-1D) were found to express c-*kit* mRNA (Figure 2).

Purification and N-Terminal Sequence Analysis of MGF

MGF was purified from +/+ conditioned medium as described in Experimental Procedures. A summary of the purification protocol is shown in Table 1. MGF is extremely hydrophobic, eluting from a C_4 reverse-phase column at

Table 1. Purification Summary

Fraction	Volume	Protein	Units	Specific Activity (U/mg)	Yield (% units)	Purification Fold
Conditioned medium	39 liters	6.25 g[a]	76,147	12	100	1
S-Sepharose + DEAE–Sephacel	345 ml	2.46 g[a]	78,786	32	103	2.7
Phenyl–Sepharose + YM10	20 ml	204 mg[b]	64,914	318	85	27
Matrex Gel Blue A	118 ml	33.4 mg[b]	37,389	1,118	49	93
Vydac C_4 reverse-phase HPLC	3.6 ml	150 μg[b]	34,257	228,380	45	19,032
Aquapore butyl reverse-phase HPLC	3.2 ml	42 μg[b]	20,905	497,738	27	41,478

[a] Protein determined by Bio-Rad (Richmond, CA) protein assay with BSA as standard.

[b] Protein estimated from silver-stained SDS–PAGE gels; therefore, specific activity values are approximate.

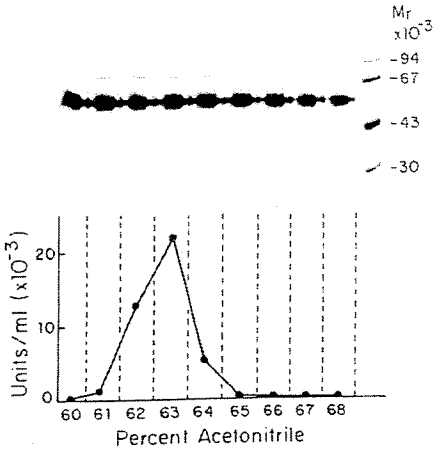

Figure 3. SDS–PAGE Analysis of MGF Purified through S-Sepharose, DEAE–Sephacel, Phenyl–Sepharose, Matrex Gel Blue A, and Reverse-Phase HPLC with a Vydac C_4 Column

The silver-stained gel is shown; directly under each lane, MGF units per ml (determined in the MC-6 bioassay) are indicated (y axis) versus the percentage acetonitrile (x axis). Molecular weight markers are shown in the rightmost lane.

62%–65% acetonitrile. Bioactivity, as assessed by the ability to stimulate the proliferation of MC-6 cells, was shown to correlate with the presence of a 31 kd protein on silver-stained gels of proteins from the HPLC fractions (Figure 3). This is the approximate molecular weight of MGF previously estimated by gel filtration (Boswell et al., 1990).

To demonstrate that the 31 kd protein was MGF, +/+ conditioned medium was purified through S-Sepharose, DEAE–Sephacel, phenyl–Sepharose, Matrex Gel Blue A, reverse-phase HPLC on a C_4 Vydac column, and reverse-phase HPLC on an Aquapore butyl column as described in Experimental Procedures. Fractions containing MGF were pooled and radiolabeled with ^{125}I. Comparison of silver-stained SDS–PAGE gels of this MGF preparation and autoradiography of the labeled material demonstrated that 31 and 36 kd protein species were labeled with ^{125}I (Figure 4A). Longer exposures of ^{125}I-labeled material did not reveal any additional protein species in the 20–40 kd molecular weight range (data not shown). The 31 kd protein specifically bound to MC-6 cells (MGF responder) but not to 32D cells (MGF nonresponder) (Figure 4B). This binding to MC-6 cells was inhibited by the addition of a 20-fold molar excess of unlabeled MGF obtained from the same fraction that was used for radiolabeling. The 36 kd protein also demonstrated the same pattern of specific, competable binding to MC-6 cells but not to 32D cells (Figure 4B).

Purified MGF was transferred to a polyvinylidene difluoride (PVDF) membrane for N-terminal sequence analysis. Silver-stained gels of this purified material indicated that the active fractions after Aquapore HPLC purification contained diffusely staining proteins of 31 and 36 kd (which were the same as those identified in the cell binding studies) as well as three larger, contaminating proteins of greater than 55 kd, which showed only nonspecific binding to both MC-6 and 32D cells (Figure 4B). Both the 31 and 36 kd proteins were transferred to PVDF membrane for sequencing and showed identical N-termini (Table 2). Searches of the GenBank, Swiss-Prot, and NBRF protein data bases indicated that this sequence was unique.

Identification of MGF As a Ligand for c-*kit*

To determine whether MGF might specifically bind to c-*kit*, ^{125}I-labeled MGF was added to MC-6 and 32D cells in the absence and presence of excess unlabeled MGF and treated with the cross-linking reagent bis(sulfosuccinimidyl) suberate (BS3). Cell lysates were then prepared and immunoprecipitated with a polyclonal antiserum that recognizes the C-terminal portion of c-*kit*. The immunoprecipitated proteins were separated by SDS–PAGE and analyzed by autoradiography. A cross-linked complex was specifically immunoprecipitated by the anti-c-*kit* antiserum from MC-6 cell lysates incubated in the absence of unlabeled competitor; this complex had an approximate molecular weight of 175–180 kd (Figure 4C). Preimmune sera did not immunoprecipitate this complex, nor was the complex observed when cross-linking was done in the presence of excess unlabeled MGF. Subtracting the molecular weight of the c-*kit* protein (145 kd) from the weight of the cross-linked complex indicated that the ligand is about 30–35 kd. MGF is the only detectable protein in the bioactive HPLC fractions in this molecular weight range and represents the only protein species showing specific, unlabeled competable binding to MC-6 cells. No evidence was seen for MGF binding to 32D cells, which do not express c-*kit* mRNA.

We also examined the binding and cross-linking of iodinated MGF to COS-7 cells that had been transfected with a cDNA encoding the murine c-*kit* gene. After binding and cross-linking of the radiolabeled material to the cells, lysates were prepared and incubated with antibodies to the C-terminus of the c-*kit* protein. As with the MC-6 cells, a cross-linked complex of approximately 180 kd was specifically immunoprecipitated by the anti-c-*kit* antiserum; no such complex was precipitated with this antiserum from nontransfected COS-7 cells, nor from the c-*kit*-transfected cells using preimmune serum (Figure 4D). An identical complex was immunoprecipitated from the c-*kit*-transfected COS-7 cells with an antiserum to the kinase domain of the c-*kit* protein (data not shown). Finally, the addition of excess unlabeled MGF blocked the formation of the radiolabeled complex.

Antisera specificity in these experiments was confirmed by immunoprecipitation studies using [^{35}S]methionine-labeled MC-6, 32D, and Rat-2 cells stably expressing a c-*kit* cDNA clone (data not shown). The antiserum to the C-terminus of c-*kit* specifically immunoprecipitated a protein of 145–150 kd from lysates of c-*kit*-transfected Rat-2 cells or MC-6 cells, but not from nontransfected Rat-2 cells or 32D cells (data not shown). Preimmune sera did not immunoprecipitate any detectable c-*kit* proteins from any of the aforementioned cells. Two additional antisera raised against synthesized peptides corresponding to amino acids 76–89 and 730–743 of the c-*kit*-encoded protein

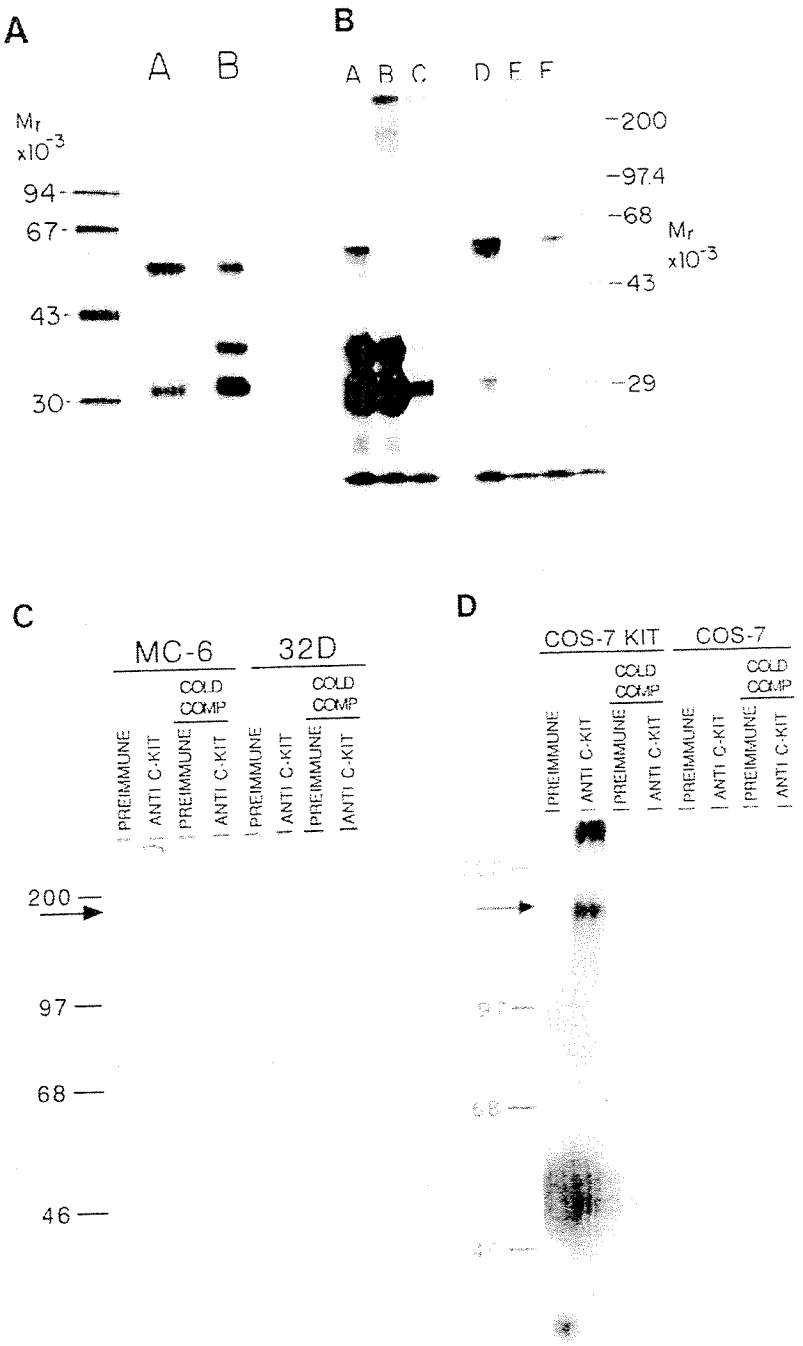

Figure 4. Binding of ^{125}I-MGF to MGF-Responsive and MGF-Nonresponsive Cells, and Cross-Linking Data Identifying MGF As a Ligand for c-*kit*

(A) Comparison of a silver-stained SDS–PAGE gel of purified MGF (from Table 1) in lane A with an autoradiograph of the same MGF preparation (after radiolabeling with ^{125}I) in lane B.

(B) Results of whole-cell binding and cross-linking studies that were done on MGF-responsive MC-6 cells (lanes A–C) and MGF-nonresponsive 32D cells (lanes D–F). In these experiments, ^{125}I-MGF was added to cells in the absence or presence of an excess of nonlabeled MGF for 30 min at 37°C. After extensive washing of the cells to remove unbound ^{125}I-MGF, cell lysates were prepared and proteins were separated by SDS–PAGE for analysis by autoradiography. In some cases, the cross-linking agent BS3 was added to cells after they were washed. Lanes A and D show the radiolabeled proteins bound to MC-6 and 32D cells, respectively, in the absence of cross-linker. Lanes B and E show proteins bound to cells in the presence of BS3, and lanes C and F show proteins bound in the presence of both BS3 and excess nonlabeled MGF.

(C and D) Results of cross-linking and immunoprecipitation experiments using antiserum to the C-terminus of the c-*kit*-encoded protein. COLD COMP refers to cross-linking done in the presence of excess nonlabeled MGF. In (C), ^{125}I-MGF was cross-linked to MC-6 and 32D cells (see Experimental Procedures); cell lysates were then prepared, incubated with either preimmune serum or anti-c-*kit* antiserum, and then immunoprecipitated and autoradiographed. In (D), ^{125}I-MGF was cross-linked to COS-7 cells that had been transfected with either a murine c-*kit* cDNA (COS-7 KIT) or vector only (COS-7) as described in Experimental Procedures; cell lysates were then made and incubated with either preimmune serum or anti-c-*kit* antiserum, followed by immunoprecipitation and autoradiography. Arrows in (C) and (D) show the position of the cross-linked ^{125}I-MGF–c-*kit* complex.

showed a pattern of reactivity identical to that of the antiserum used in the present studies (data not shown).

In addition to the 175–180 kd complex immunoprecipitated by the anti-c-*kit* antiserum, a much higher molecular weight complex was precipitated in the MC-6 and transfected COS-7 cells (Figures 4C and 4D); this band may result from the cross-linking of ligand to receptor dimers. Evidence for ligand-induced receptor dimerization (or oligomerization) has been seen with other members of the receptor tyrosine kinase family, including the epidermal growth factor receptor (Cochet et al., 1988) and the platelet-derived growth factor receptor (Andrius Kazlauskas, personal communication).

Discussion

The pleiotropic nature of mutant alleles at the steel (*Sl*) and dominant spotting (*W*) loci has been intensely studied and has been of particular interest in the study of hematopoiesis (Russell, 1979). The defect occurs at the level of

Table 2. N-Terminal Amino Acid Sequence of the 31 and 36 kd Forms of MGF

	1	2	3	4	5	6	7	8	9	10	11	12
31 kd	K	E	I	(C)	G	N	P	V	T	D	N	V
36 kd	K	E	I	(C)	G	N	P	V	T	D	N	V
	13	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
31 kd	K	D	I	T	K	L	V	A	N	L	P	N
36 kd	K	D	I	T	K	L	V	A	N	L	P	N
	25	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**
31 kd	D	Y	M	I	T	L	N	Y	V	A	G	M
36 kd	D	Y	M	I	T	L	N	Y	V	A	G	M
	37	**38**	**39**	**40**	**41**	**41**	**43**	**44**	**45**	**46**	**47**	**48**
31 kd	D	V	L	P	S			W		Y	D	M
36 kd	D	V	L	P	S	W		(W)	L	(D)	(D)	(M)
	49	**50**	**51**	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**
31 kd	V	I	Q			(W)						
36 kd	(I)	Q	L	(D)								

N-terminal amino acid sequence analysis was performed on MGF purified as described in Experimental Procedures after transfer to PVDF membrane. Both the 31 and 36 kd species, which demonstrated specific binding to MGF-responsive cells, were sequenced.

the PHSC in both mouse mutants; however, the erythroid (Russell, 1979; Harrison and Russell, 1972) and mast cell (Russell, 1979; Kitamura and Go, 1978; Kitamura et al., 1978) lineages are disproportionately affected. *W* mice possess an intrinsic abnormality at the level of the PHSC (Russell, 1979; Russell et al., 1956; Russell and Bernstein, 1968; Murphy et al., 1973). This has been demonstrated in experiments in which histocompatible, normal hematopoietic cells from bone marrow, spleen, or fetal liver were injected into W/W^v recipients with subsequent cure of the hematopoietic defects (Russell, 1979; Russell et al., 1956, Russell and Bernstein, 1968; Murphy et al., 1973; Bernstein, 1970). Similar injection of normal hematopoietic cells into *Sl* mice (Sl/Sl^d) failed to reconstitute hematopoiesis (Russell, 1979; Bernstein, 1970; Harrison and Russell, 1972; Fried et al., 1973); however, transplantation of Sl/Sl^d hematopoietic tissue did reconstitute hematopoiesis in W/W^v mice (Russell, 1979; Bernstein, 1970). This indicates that the defect in Sl/Sl^d mice is not intrinsic to PHSCs. Intergenotype transplantation of spleens has shown that the defect in hematopoiesis in *Sl* mice is related to the microenvironment (Russell, 1979; Bernstein, 1970; Fried et al., 1973). Similar conclusions have been reached in vitro using long-term bone marrow cultures (Dexter and Moore, 1977).

Recent studies have demonstrated that the c-*kit* proto-oncogene maps to the *W* locus on murine chromosome 5 (Chabot et al., 1988; Geissler et al., 1988). A number of mutant alleles have been sequenced and have shown either rearrangement, deletions, or point mutations in the c-*kit* coding sequence (Geissler et al., 1988; Nocka et al., 1990; Tan et al., 1990; Reith et al., 1990). In addition, studies of c-*kit* mRNA levels in the affected tissues of mutant *W* mice, and reductions in c-*kit* tyrosine kinase activity in the mutants, provide strong supportive evidence for the identity of c-*kit* as a product of the *W* locus (Nocka et al., 1989, 1990). The complementary nature of the *W* and *Sl* mutations led to the hypothesis that the gene product of the *Sl* locus encodes a ligand for c-*kit* (Chabot et al., 1988; Geissler et al., 1988).

We have previously described the preliminary characterization of a stromal cell clone derived from the bone marrow of a normal mouse (+/+) (Boswell et al., 1990). This cell line, but not a stromal cell clone derived from an Sl/Sl^d bone marrow culture, could support the survival of two IL-3-dependent mast cell lines (H7, MC-6) in the absence of exogenous growth factors (Boswell et al., 1990). Supernatants from +/+ cells contain a growth factor (MGF) for H7 and MC-6 cells (Boswell et al., 1990). The mast cell phenotype of the responder cell lines coupled with the inability of similar stromal cells from mutant *Sl* mice to support the survival of H-7 and MC-6 cells led us to postulate that the growth factor from +/+ cells represented a ligand for c-*kit*. This hypothesis was further strengthened by the direct correlation between responsiveness of a panel of IL-3-dependent cell lines to partially purified +/+ supernatants and the expression of c-*kit* mRNA.

Biologically active MGF was purified by a series of chromatographic separation procedures. Following the initial reverse-phase HPLC purification on a Vydac C_4 column, a correlation between bioactivity in the MC-6 proliferation assay and the presence of 31 and 36 kd protein species was observed. Whole-cell binding experiments with ^{125}I-labeled, purified MGF were carried out to identify the protein species responsible for biological activity. Specific binding of iodinated MGF, which was competed by excess unlabeled material, was observed on MC-6 cells and correlated with proteins of 31 kd and 36 kd. Sequence analysis of these two proteins indicated that they had identical N-termini. The difference in molecular weight may reflect differential glycosylation, as silver-stained gels revealed very diffuse bands for both protein species. Preliminary N-glycanase digestions of purified MGF confirm the presence of N-linked sugars (C. R., unpublished data), indicating that MGF is a glycoprotein.

Cross-linking studies using radiolabeled MGF followed by immunoprecipitation with an antiserum specific to the C-terminus of c-*kit* indicated that a ligand for c-*kit* was present in the ^{125}I-labeled fraction. The molecular weight of the observed receptor–ligand complex indicated that the ligand is 30–35 kd in size. The only detectable protein species by either silver staining of SDS–PAGE gels or autoradiography of radiolabeled material in the 30–35 kd range were the two protein species for which N-terminal sequence was obtained. This indicated that MGF was a ligand for the c-*kit* proto-oncogene. In the accompanying report (Copeland et al., 1990), isolation of cDNA clones for MGF and chromosomal localization studies indicate that the gene encoding MGF maps near the *Sl* locus on mouse chromosome 10 and is deleted in a number of *Sl* alleles.

Throughout this paper we have referred to MGF as a ligand, rather than the ligand, for the c-*kit* proto-oncogene. We cannot rule out the possiblity that there are multiple ligands capable of binding to c-*kit*. For example, it has been shown that GM-CSF and IL-3 bind to a common cell surface receptor (Park et al., 1989); these cytokines are

closely linked on human chromosome 5 (Le Beau et al., 1987; Huebner et al., 1985). Thus, we recognize that ligands in addition to MGF may be capable of binding to the c-*kit*-encoded receptor.

Experimental Procedures

Cell Lines

The NFS.N1.H7 (H7), NFS.N1.MC6 (MC-6), and +/+ cell lines have been described (Boswell et al., 1990). H7 and MC-6 cells were grown in Click's (Altick Associates, River Falls, WI) or McCoy's 5A (GIBCO, Grand Island, NY) medium, respectively, supplemented with 10% heat-inactivated (56°C, 30 min) fetal bovine serum (FBS; Flow Laboratories, McLean, VA), 5×10^{-5} M 2-mercaptoethanol (GIBCO), and recombinant murine IL-3 (Park et al., 1986a). The FDC-P2, FDC-P2-1D, DA-1, and 32D cell lines have also been described (Williams et al., 1990), and these lines were grown in RPMI 1640 (GIBCO) with 10% FBS, 5×10^{-5} M 2-mercaptoethanol, and 10% WEHI-3B conditioned medium as a source of IL-3 (Prestige et al., 1984), with the exception of FDC-P2-1D, which was grown in 20 ng/ml of purified recombinant murine GM-CSF (Park et al., 1986b). The +/+ cell line was routinely passaged twice weekly in McCoy's medium with 10% FBS. For collection of conditioned medium, cells were grown in roller bottles (Corning, Corning, NY) until they reached approximately 80% confluency. Medium was then removed, and cells were refed with Opti-MEM (GIBCO) serum-free medium and incubated in a 37°C warm room for 5 days.

Proliferation Assays

Cell proliferation was assessed in [^{3}H]thymidine incorporation assays. Serial dilutions of growth factors were made in 96-well flat-bottom tissue culture plates (Costar, Cambridge, MA) in the appropriate growth medium at a final volume of 50 μl. Cells were seeded at time zero (50 μl volume), and the plates were incubated at 37°C in a fully humidified atmosphere of 6.5% CO_2 in air for 24 hr (40 hr for DA-1). Cells were pulsed with 0.5 μCi per well of [^{3}H]thymidine (25 Ci/mmol, Amersham, Arlington Heights, IL) and incubated for an additional 5–6 hr. Samples were harvested onto glass fiber filters and counted by liquid scintillation spectrometry.

Purification of MGF

The conditioned medium from +/+ cells was acidified to pH 2.75 by the addition of concentrated HCl and buffered with sodium citrate (pH 3.25) to a final concentration of 5 mM. The acidified conditioned medium was applied to an S-Sepharose Fast Flow column (Pharmacia, Piscataway, NY) equilibrated with 0.1 M NaCl, 5 mM sodium citrate (pH 3.25) and eluted with 0.1 M EPPS (pH 8.5) (Sigma Chemical, St. Louis, MO). The eluted protein peak was diluted 1:3 with water and applied to a DEAE-Sephacel (Pharmacia) column equilibrated with 10 mM EPPS (pH 8.5). The column was eluted with an NaCl gradient (0–0.5 M NaCl in 10 mM EPPS [pH 8.5]). Active fractions were pooled following bioassay, adjusted to 2 M ammonium sulfate (ultra pure; ICN Biomedicals, Cleveland, OH), 20 mM Tris-HCl (pH 7.5) (Sigma), centrifuged to remove precipitated proteins, and applied to a phenyl-Sepharose CL-4B (Pharmacia) column equilibrated with 2 M ammonium sulfate, 20 mM Tris-HCl (pH 7.5). The column was washed and eluted with a reverse gradient from 2 M to 0.5 M ammonium sulfate in 20 mM Tris-HCl (pH 7.5). Aliquots of fractions were dialyzed against PBS (6000–8000 dalton cutoff membrane; Bethesda Research Laboratories, Gaithersburg, MD) and bioassayed. Active fractions were pooled, and buffer was exchanged with 5 mM $CaCl_2$, 20 mM HEPES (pH 7.5) and concentrated 10–12 times with an Amicon (Amicon Corp., Lexington, MA) stirred cell concentrator with a YM10 membrane. The concentrated phenyl-Sepharose pool was passed over a Matrex Gel Blue A (Amicon Corp.) column washed with 5 mM $CaCl_2$, 20 mM HEPES (pH 7.5). The Blue A pool was acidified with trifluoroacetate (TFA) to 0.1% (v/v), 0.22 μm filtered (cellulose acetate filter), and chromatographed by HPLC on a Radial Pak (Waters) 0.8 × 10 cm column packed with Vydac C_4 (15–20 μm; Vydac, Hesperia, CA) equilibrated in 0.1% TFA in water. The column was eluted with a 1% per min linear gradient of acetonitrile (0.1% [v/v] TFA) at a flow rate of 0.8 ml/min.

Fractions with MGF activity were dried to 10% of the original volume with a Speed Vac apparatus (Savant Instruments) and rechromatographed by HPLC on an Aquapore butyl column (7 μm, 30 × 2.1 mm; Applied Biosystems, Foster City, CA) preequilibrated with 0.1% TFA in H_2O at a flow rate of 0.2 ml/min.

Bioactive material from the Aquapore column was dried down and resuspended in gel sample buffer and heated to 110°C for 5 min. The protein was then separated using a 10% Laemmli gel and transferred onto PVDF membrane (Pro-Blot, Applied Biosystems) using constant current (60 V setting) for 1 hr. The proteins were visualized by staining the PVDF membrane for 5 min with Coomassie blue at 0.1% in 10% acetic acid, 50% HPLC-grade methanol, and HPLC-grade H_2O. The membrane was destained with the above solution without Coomassie blue and then washed extensively with HPLC-grade H_2O before drying. Stained proteins were excised with a single-edged razor and loaded onto a model 477 protein sequencer (Applied Biosystems) with an on-line model 120A HPLC analyzer (Applied Biosystems) for N-terminal sequencing.

Radiolabeling of MGF and Binding to Intact Cells

Labeling of MGF with ^{125}I was performed with a previously described Enzymobead procedure (Park et al., 1986b) using aliquots of MGF active fractions (Vydac C_4 or Aquapore derived) containing 1–5 μg of total protein diluted to a final volume of 50 μl in 0.2 M $NaPO_4$ (pH 7.2). For a cold competition stock, an aliquot of the HPLC fraction used for radiolabeling was exchanged into an equivalent volume of binding medium (RPMI 1640, 20 mM HEPES, 0.2% NaN_3, 2.5% BSA [pH 7.2]) by evaporation under nitrogen. Whole-cell binding studies were carried out essentially as previously described (Namen et al., 1988b). In brief, 6×10^6 MC-6 (MGF responder) or 32D (MGF nonresponder) cells were combined with 3×10^7 cpm of labeled MGF with or without 50 μl of unlabeled competition stock in a final volume of 150 μl of binding medium. Tubes were incubated for 30 min at 37°C with rocking, and the cells were then washed by centrifugation three times in cold PBS. After the final wash, cell pellets were extracted for 20 min at 4°C in 50 μl of PBS containing 1% Triton X-100, PMSF (2 mM), pepstatin (10 μM), leupeptin (10 μM), o-phenathroline (2 mM), 0.4 mg/ml trypsin inhibitor, 250 mM benzamidine, 50 mM EDTA, 1 mg/ml E-64, 0.02% NaN_3, and EGTA (2 mM) (Sigma). The suspension was centrifuged, and supernatants were analyzed by autoradiography after SDS-PAGE analysis as described (Park et al., 1986b). For affinity cross-linking experiments, cells were incubated for 30 min at 25°C with BS3 (Pierce Chemical, Chicago, IL) at a final concentration of 0.1 mg/ml prior to extraction and SDS-PAGE analysis as described (Park et al., 1986b).

Cloning of the Murine c-*kit* cDNA

A cDNA corresponding to the murine c-*kit* gene was cloned using the polymerase chain reaction (Saiki et al., 1985) from poly(A)$^+$-selected mRNA isolated from MC-6 cells. The 3.0 kb cDNA was sequenced and compared with the published murine c-*kit* sequence. Two changes at the amino acid level were noted upon comparison with the published sequence (Qiu et al., 1988): our clone had an alanine instead of glutamic acid at amino acid 207, and a glycine instead of an alanine at amino acid 777. Multiple cDNA clones were sequenced to ensure that these changes were not random artifacts of the PCR amplification technique. This c-*kit* cDNA was inserted into the pDC302 expression vector plasmid for transient expression in COS-7 cells as previously described (Mosley et al., 1989).

Antiserum to Murine c-*kit*-Encoded Protein and Immunoprecipitation Studies

Anti-peptide antiserum to the murine c-*kit* protein was generated as follows. A C-terminal peptide corresponding to the last 10 amino acids of the published sequence was synthesized, conjugated with ovalbumin, and injected into rabbits. Serum used for immunoprecipitation experiments was obtained from the rabbits after an initial injection and four subsequent boosts. Metabolic labeling of cells was performed with [^{35}S]methionine (Amersham) as previously described (Namen et al., 1988a). Immunoprecipitations were performed on lysates of ^{35}S-labeled cells, or from cell lysates obtained after cross-linking as described under "Radiolabeling of MGF and Binding to Intact Cells." Five microliters of a 1:10 dilution of either preimmune serum or antiserum raised against the C-terminal c-*kit* peptide was added to 100 μl of cell lysate and placed at 4°C for 1 hr. Samples were then incubated with protein A-Sepharose (Pharmacia) for 30 min, and the immune com-

plexes were washed extensively prior to SDS–PAGE and autoradiography as previously described (Lyman and Rohrschneider, 1989).

Northern Blot Analysis of c-*kit* mRNA

Two micrograms of poly(A)$^+$-selected mRNA isolated from a variety of cell lines was electrophoresed in 1.2% agarose gels in the presence of 0.74% formaldehyde (Maniatis et al., 1982) and analyzed by Northern hybridization. After electrophoresis, the gel was blotted onto Hybond-N (Amersham) and mRNA was transferred by capillary action in 10× SSC (3 M NaCl, 0.3 M sodium citrate). mRNA was subsequently cross-linked to the filter by exposure to UV light. A 21 nucleotide antisense probe (5′-TGTACCGTCACGCTGAGCAGC-3′) corresponding to a portion of the extracellular domain of the c-*kit* gene was radiolabeled with [γ-^{32}P]ATP using T4 polynucleotide kinase and hybridized to the nitrocellulose blot. Filters were incubated overnight at 55°C in hybridization buffer (60 mm Tris [pH 8.0], 2 mM EDTA, 5× Denhardt's solution, 6× SSC, 0.1% lauryl sarcosine, 0.5% NP40, 200 μg/ml sheared salmon sperm DNA) with the radiolabeled probe and then washed for 30 min in 0.1× SSC, 0.1% SDS at 55°C. Autoradiography was performed at −70°C using Kodak X-Omat AR film (Kodak, Rochester, NY).

Acknowledgments

The authors wish to thank the skilled cell culturists of the Immunex Department of Bioresources for production of conditioned medium and the expert technical assistance of Vickie Craig, Felicia Koerner, Christy McNutt, Laura James, and Dauphine Torrance. We also thank Janet Merriam, Paul Sleath, and Carl March for synthesizing the c-*kit* peptides, JoDee Zappone and Ric Raunio for preparation of the anti-c-*kit* antisera, and Linda Troup for preparation of the manuscript.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 USC Section 1734 solely to indicate this fact.

Received July 30, 1990; revised August 9, 1990.

References

Bernstein, S. E. (1970). Tissue transplantation as an analytic and therapeutic tool in hereditary anemias. Am. J. Surg. *119*, 448–451.

Bernstein, S. E., Russell, E. S., and Keighley, G. H. (1968). Two hereditary mouse anemias (*Sl/Sl^d* and *W/W^v*) deficient in response to erythropoietin. Ann. NY Acad. Sci. *149*, 475–485.

Boswell, H. S., Albrecht, P. R., Shupe, R. E., Williams, D. E., and Burgess, J. (1987). Role of stromal populations in hemopoietic stem cell proliferation. I. Physically distinct subpopulations of hemopoietic stem cells and stromal progenitors determine long-term culture hemopoiesis. Exp. Hematol. *15*, 46–53.

Boswell, H. S., Mochizuki, D. Y., Burgess, G. S., Gillis, S., Walker, E. B., Anderson, D., and Williams, D. E. (1990). A novel mast cell growth factor (MCGF-3) produced by marrow-adherent cells that synergizes with interleukin-3 and interleukin-4. Exp. Hematol. *18*, 794–800.

Chabot, B., Stephenson, D. A., Chapman, V. M., Besmer, P., and Bernstein, A. (1988). The proto-oncogene c-*kit* encoding a transmembrane tyrosine kinase receptor maps to the mouse *W* locus. Nature *335*, 88–89.

Cochet, C., Kashles, O., Chambaz, E. M., Borrello, I., King, C. R., and Schlessinger, J. (1988). Demonstration of epidermal growth factor–induced receptor dimerization in living cells using a chemical covalent cross-linking agent. J. Biol. Chem. *263*, 3290–3295.

Copeland, N. G., Gilbert, D. J., Cho, B. C., Donovan, P. J., Jenkins, N. A., Cosman, D., Anderson, D., Lyman, S. D., and Williams, D. E. (1990). Mast cell growth factor maps near the steel locus on mouse chromosome 10 and is structurally altered in a number of steel alleles. Cell *63*, this issue.

Dexter, T. M., and Moore, M. A. S. (1977). *In vitro* duplication and "cure" of hematopoietic defects in genetically anemic mice. Nature *269*, 412–414.

Dexter, T. M., Allen, T. D., and Lajtha, L. G. (1971). Conditions controlling the proliferation of hemopoietic stem cells in vitro. J. Cell. Physiol. *91*, 335–344.

Eliason, J. F., Thorens, B., Kindler, V., and Vassili, P. (1988). The roles of granulocyte macrophage colony-stimulating factor and interleukin-3 in stromal cell mediated hemopoiesis in vivo. Exp. Hematol. *16*, 307–312.

Fried, W., Chamberlin, W., Knopse, W. H., Husseini, S., and Trobaugh, F. E., Jr. (1973). Studies on the defective hematopoietic microenvironment of *Sl/Sl^d* mice. Br. J. Haematol. *24*, 643–650.

Geissler, E. N., Ryan, M. A., and Housman, D. E. (1988). The dominant-white spotting (*W*) locus of the mouse encodes the c-*kit* proto-oncogene. Cell *55*, 185–192.

Gualteri, R. J., Yong, G., Shadduck, R. K., Waheed, A., and Banks, A. (1987). Identiification of the hematopoietic growth factors elaborated by bone marrow stromal cells using antibody neutralization analysis. Exp. Hematol. *15*, 883–889.

Harrison, D. E., and Russell, E. S. (1972). The response of *W/W^v* and *Sl/Sl^d* anemic mice to haematopoietic stimuli. Br. J. Haematol. *22*, 155–168.

Huebner, K., Isobe, M., Croce, C. M., Golde, D. W., Kaufman, S. E., and Gasson, J. C. (1985). The human gene encoding GM-CSF is at 5q21–q32, the chromosome region deleted in the 5q$^-$ anomaly. Science *230*, 1282–1285.

Kitamura, Y., and Go, G. (1978). Decreased production of mast cells in *Sl/Sl^d* anemic mice. Blood *53*, 492–497.

Kitamura, Y., Go, S., and Hatanaka, K. (1978). Decrease of mast cells in *W/W^v* mice and their increase by bone marrow transplantation. Blood *52*, 447–452.

Le Beau, M. M., Epstein, N. D., O'Brien, S. J., Nienhuis, A. W., Yang, Y.-C., Clark, S. C., and Rowley, J. D. (1987). IL-3 maps to human chromosome 5 and is deleted in myeloid leukemias with a del(5q). Proc. Natl. Acad. Sci. USA *84*, 5913–5917.

Li, C. L., and Johnson, G. R. (1985). Stimulation of multipotential, erythroid, and other murine hematopoietic progenitor cells by adherent cell lines in the absence of detectable multi-CSF (IL-3). Nature *316*, 633–636.

Lyman, S. D., and Rohrschneider, L. R. (1989). The kinase activity of the v-*fms* encoded protein has a low pH optimum. Oncogene Res. *4*, 149–155.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Mayer, T. C., and Green, T. C. (1968). An experimental analysis of the pigment defect caused by mutations at the *W* and *Sl* loci in mice. Dev. Biol. *18*, 62–75.

Mintz, B., and Russell, E. S. (1957). Gene induced embryological modifications of primordial germ cells in the mouse. J. Exp. Zool. *134*, 207–237.

Mosley, B., Beckmann, M. P., March, C. J., Idzerda, R. L., Gimpel, S. D., VandenBos, T., Friend, D., Alpert, A., Anderson, D., Jackson, J., Wignall, J. M., Smith, C., Gallis, B., Sims, J. E., Urdal, D., Widmer, M. B., Cosman, D., and Park, L. S. (1989). The murine interleukin-4 receptor: molecular cloning and characterization of secreted and membrane bound forms. Cell *59*, 335–348.

Murphy, E. D., Harrison, D. E., and Roths, J. B. (1973). Giant granules of beige mice, a quantitative marker for granulocytes in bone marrow transplantation. Transplantation *15*, 526–530.

Namen, A. E., Lupton, S., Hjerrild, K., Wignall, J., Mochizuki, D. Y., Schmierer, A., Mosley, B., March, C. J., Urdal, D. L., Gillis, S., Cosman, D., and Goodwin, R. G. (1988a). Stimulation of B-cell progenitors by cloned murine interleukin-7. Nature *333*, 571–573.

Namen, A. E., Schmierer, A. E., March, C. J., Overell, R. W., Park, L. S., Urdal, D. L., and Mochizuki, D. Y. (1988b). B cell precursor growth promoting activity: purification and characterization of a growth factor active on lymphocyte precursors. J. Exp. Med. *167*, 988–1002.

Nocka, K., Mujumder, S., Chabot, B., Ray, P., Cervonne, M., Bernstein, A., and Besmer, P. (1989). Expression of c-*kit* gene products in known cellular targets of *W* mutations in normal and *W* mutant mice—evi-

J. W. (1989) *J. Biol. Rhythms* 4, 201–216
3 Njus, D., Gooch, V. and Hastings, J. W. (1981) *Cell Biophys.* 3, 223–231
4 Johnson, C. H., Roeber, J. and Hastings, J. W. (1984) *Science* 223, 1428–1430
5 Moore-Ede, M. C., Sulzman, F. and Fuller, C. A. (1982) *The Clocks That Time Us*, Harvard University Press
6 Turek, F. N. (1985) *Annu. Rev. Physiol.* 47, 49–64
7 Young, M. W., Bargiello, T. A., Baylies, M. K. and Saez, L. (1989) in *Neuronal and Cellular Oscillators* (Jacklet, J. W., ed.), pp. 529–542, Marcel Dekker
8 Jacklet, J. W. (1989) in *Neuronal and Cellular Oscillators* (Jacklet, J. W., ed.), pp. 483–527, Marcel Dekker
9 McClung, C., Fox, B. A. and Dunlap, J. C. (1989) *Nature* 339, 558–562
10 Sweeney, B. M. (1987) *Rhythmic Phenomena in Plants*, Academic Press
11 Morse, D., Milos, P. M., Roux, E. and Hastings, J. W. (1989) *Proc. Natl Acad. Sci. USA* 86, 172–176
12 Nicolas, M-T., Nicolas, G., Johnson, C. H., Bassot, J-M. and Hastings, J. W. (1987) *J. Cell Biol.* 105, 723–735
13 Johnson, C. H., Inoué, S., Flint, A. and Hastings, J. W. (1985) *J. Cell Biol.* 100, 1435–1446
14 Hastings, J. W. (1982) in *Oxygenases and Oxygen Metabolism* (Nozaki, M., Yamamoto, S., Ishimura, Y., Coon, M., Ernster, L. and Estabrook, R., eds), pp. 225–237, Academic Press
15 Hastings, J. W. (1983) *J. Mol. Evol.* 19, 309–321
16 Nakamura, H., Kishi, Y., Shimomura, O., Morse, D. and Hastings, J. W. (1989) *J. Am. Chem. Soc.* 111, 7607–7611
17 Dunlap, J. and Hastings, J. W. (1981) *J. Biol. Chem.* 256, 10509–10518
18 Morse, D., Pappenheimer, A. M. and Hastings, J. W. (1989) *J. Biol. Chem.* 264, 11822–11826
19 Fritz, L., Morse, D. and Hastings, J. W. (1989) *J. Cell Sci.* 95, 321–328
20 Loros, J. J., Denome, S. A. and Dunlap, J. (1989) *Science* 243, 385–388
21 Hardin, P. E., Hall, J. C. and Rosbash, M. (1990) *Nature* 343, 536–540
22 Kloppstech, K. (1985) *Planta* 165, 502–506
23 Nagy, F., Kay, S. and Chua, N-H. (1988) *Genes Devel.* 2, 376–382
24 Robinson, B. G., Frim, D. M., Schwartz, W. J. and Majzoub, J. A. (1988) *Science* 241, 342–344
25 Milos, P., Morse, D. and Hastings, J. W. (1989) *Naturwissenschaften* 77, 87–89
26 Kozak, M. (1988) *J. Cell Biol.* 107, 1–7
27 Kearney, K. R. and Nomura, M. (1987) *Mol. Gen. Genet.* 210, 60–68
28 Freedman, L. P., Zengel, J. M., Archer, R. H. and Lindahl, L. (1987) *Proc. Natl Acad. Sci. USA* 84, 6516–6520
29 Hentze, M. W., Caughman, S. W., Roualt, T. A., Barriocanal, J. G., Dancis, A., Harford, J. B. and Klausner, R. D. (1987) *Science* 238, 1570–1573
30 Taylor, W. R., Dunlap, J. C. and Hastings, J. W. (1982) *J. Exp. Biol.* 97, 121–136
31 Karakashian, M. W. and Hastings, J. W. (1962) *Proc. Natl Acad. Sci. USA* 48, 2130–2137
32 Broda, H., Brugge, D., Homma, K. and Hastings, J. W. (1985) *Cell Biophys.* 8, 47–67
33 Broda, H., Gooch, V. D., Taylor, W., Aiuto, N. and Hastings, J. W. (1986) *J. Biol. Rhythms* 1, 251–263
34 Homma, K. and Hastings, J. W. (1989) *Exp. Cell Res.* 182, 635–644
35 Hastings, J. W. and Sweeney, B. M. (1957) *Proc. Natl Acad. Sci. USA* 43, 804–811
36 Roenneberg, T. and Hastings, J. W. (1988) *Naturwissenschaften* 75, 206–207
37 Roenneberg, T., Nakamura, H. and Hastings, J. W. (1988) *Nature* 334, 432–434

The amino acid sequences of several, recently cloned cytokine receptors show significant homologies, primarily in their extracellular, ligand-binding domains. With one exception, their cognate cytokines mediate biological activities on a variety of hematopoietic cell types; thus we have designated the receptors as the hematopoietin receptor superfamily.

CYTOKINE RESEARCH has been a field of extraordinary growth during the past few years. Although the nomenclature of this field has been fragmented into growth factors, hormones, lymphokines and colony stimulating factors (reflecting the different contexts in which the cytokines were discovered), there has been a general realization that these subcategories are rather artificial, and that most cytokines exhibit a range of biological activities on disparate cell types. In order to understand how these diverse effects are mediated in the responder cell, many groups have begun to characterize cytokine receptors at the molecular level. As a result of these efforts, it has become clear that cytokine receptors can be grouped into a number of families on the basis of structural homologies. For example, the receptors for interleukin-1 (IL-1) and platelet-derived growth factor are members of the immunoglobulin superfamily, based on homologies between their extracellular, ligand-binding domains; this superfamily includes cell-surface adhesion molecules, T-cell antigen receptors and other membrane glycoproteins[1]. Recently, a new cytokine receptor superfamily has been recognized as more of its members have been cloned and sequenced. This article will review the structural characteristics of this new family of receptors, which we have termed the hematopoietin receptor superfamily, and will speculate on the possible function of these elements.

D. Cosman, S. D. Lyman, R. L. Idzerda, M. P. Beckmann, L. S. Park, R. G. Goodwin and **C. J. March** are at the Immunex Corporation, 51 University Street, Seattle, WA 98101, USA.

The hematopoietin receptor superfamily

The members of this family include the receptors for: IL-2 (the β-subunit, also known as p75), IL-3, IL-4, IL-6, IL-7, granulocyte-macrophage colony stimulating factor, prolactin, growth hormone and erythropoietin. A complete description of the biological activities of the ligands for these receptors is beyond the scope of this article but can be found elsewhere[2,3]. In brief, IL-2 stimulates proliferation of T cells, and IL-3 acts to promote growth of a variety of hematopoietic lineages, including early stem cells. Interleukin-4 is also a hematopoietic growth and differentiation factor with effects on T, B and mast cells; IL-6 has even wider activities as a growth and differentiation factor for T cells, B cells, hepatocytes, and neuronal cells (among others). Interleukin-7 is a growth factor for pre-B cells and both immature and mature T cells; GM-CSF is a growth and differentiation factor for neutrophils, basophils, eosinophils, and monocytes/macrophage.

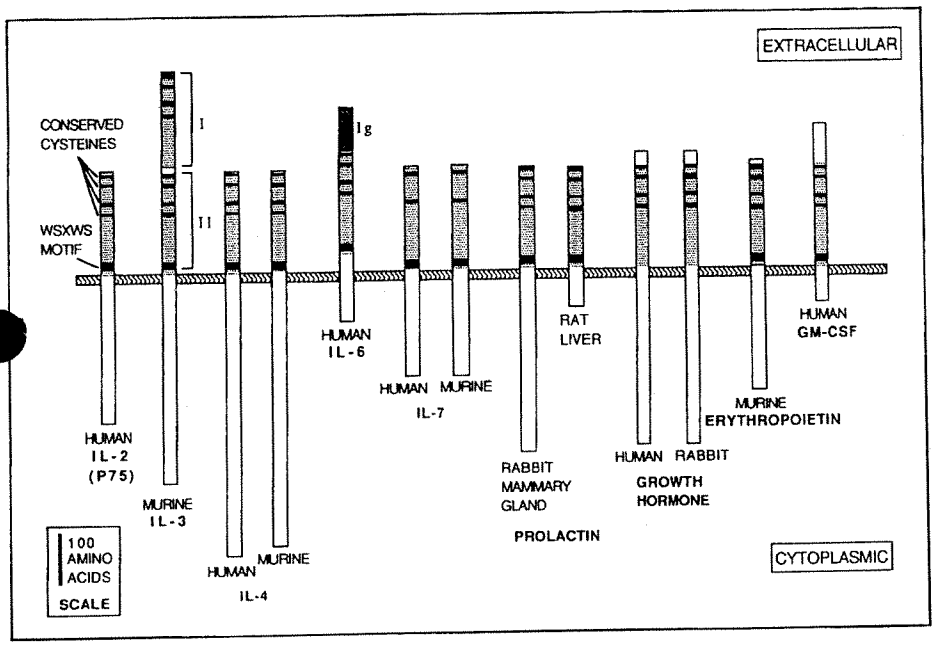

Figure 1

The hematopoietin receptor superfamily. Schematic representations of the structures of all the known members of the family are shown. Horizontal bars represent conserved cysteine residues. The black boxes represent the conserved Trp-Ser-X-Trp-Ser (WSXWS) motif. The stippled areas show the stretch of ~210 amino acids within which the receptor homologies are contained. The immunoglobulin-like domain at the N-terminus of the IL-6 receptor is also indicated.

Prolactin is best characterized for promoting the development of the mammary gland and lactation, but has also been shown to stimulate lymphocyte proliferation[4]. Growth hormone induces expression of the insulin-like growth factor-1 from the liver (and through this secondary action promotes growth)[5], and erythropoietin stimulates growth and differentiation of immature erythroblasts[6].

The receptors for growth hormone (from rabbit and human liver) and prolactin (from rat liver and rabbit mammary gland) were the first members of this family to be cloned[7–9]. These receptor sequences showed considerable homology, which was not unexpected since growth hormone and prolactin are related and can even bind to each other's receptor. The structure of these receptors was recognized as prototypical of a new class of transmembrane receptors. The IL-6 receptor (IL-6R) was cloned shortly thereafter from a human natural killer (NK) cell line[10]. It was found to have a single, N-terminal, immunoglobulin-like domain, but the structural relationships between the remainder of the extracellular domain of the IL-6 receptor and the prolactin and growth hormone receptors was not recognized at the time because of the small number of conserved amino acids. The next receptors cloned in this family were the murine erythropoietin receptor (from an erythroleukemia cell line) and the IL-2 receptor β-subunit (from a human NK cell line)[11,12]. The considerable homology between these latter two receptor sequences was subsequently recognized[13]. The first published reports of a larger family of receptors came simultaneously from our group (as a result of our cloning of the murine and human IL-4 receptors), from Gearing *et al.* describing the human GM-CSF receptor, and by Bazan[14–17]. Subsequent cloning of the murine IL-3 and human and murine IL-7 receptors served to expand this family even further[18,19].

Structural homologies

A schematic representation of the members of the superfamily is shown in Fig. 1. All the receptors are type I membrane glycoproteins, oriented with their N-termini outside the plasma membrane as shown and with a single hydrophobic transmembrane domain. The molecules are translated with an N-terminal signal sequence that directs the translocation of the nascent polypeptide across the membrane of the endoplasmic reticulum, and which is subsequently removed.

The major region of homology between the receptors occurs in the extracellular, ligand-binding domain, and is contained within a stretch of about 210 amino acids (stippled in Fig. 1). The characteristic, highly conserved features of this region are four cysteine residues located in the N-terminal half of the region, and a Trp-Ser-X-Trp-Ser motif (WSXWS) located just outside the membrane-spanning domain (Fig. 1). These features are found in all of the receptors except the IL-7 receptor (which has only two of the conserved cysteines) and the growth hormone receptor (which lacks the WSXWS motif). This conserved extracellular region appears to have been duplicated in the IL-3 receptor; however, the N-terminal half of this duplicated region lacks the WSXWS motif.

The homologies between the receptors within this conserved extracellular region were quantitated using the ALIGN program, which compares two sequences and gives them a score according to their degree of relatedness. A score of greater than three is considered to show a significant evolutionary relationship[1]. Table I shows the pairwise ALIGN scores and (in the shaded area) the percentage amino acid identity between all the members of the hematopoietin receptor superfamily. With a few exceptions (such as the murine IL-4 receptor versus the human growth hormone receptor), the scores are all greater than three, and most are in the range of four to nine. These scores are similar to those calculated between members of the immunoglobulin superfamily. The erythropoietin receptor shows relatively high ALIGN scores with all the other members of this family and thus appears to occupy a central evolutionary position.

Although the four cysteines and the WSXWS motif are the more strikingly conserved elements of this region, other conserved amino acid residues can be found. Figure 2 shows an alignment of the whole conserved region among 13 of the receptors. Amino acids conserved in six or more of the receptors are boxed. The structural or functional importance, if any, of particular residues that appear to be conserved will have to be addressed by biophysical characterization and site-specific mutagenesis, followed by ligand-binding and signal transduction assays. It is striking, however, that not only is the amino acid sequence WSXWS conserved, but also the codons for the two serine residues of this motif are limited

(with one exception), to AGC and AGT, out of the six possible serine codons. This implies a common evolutionary origin for these receptors. The high degree of codon usage conservation of this WSXWS motif may possibly be exploited to design oligonucleotide probes for the purpose of cloning new members of this receptor family.

The WSXWS sequence motif is not found only in members of the hematopoietin receptor superfamily. A search of sequence databases and the scientific literature revealed this sequence in a small number of other proteins. However, conservation of the codons for serine is not generally observed in these other proteins. Notable among these proteins are the complement precursor proteins C7, C8 and C9, as well as thrombospondin and properdin. These proteins share three common features: (1) a number of internal sequence repeats, (2) localization at, or association with the cell membrane, and (3) interaction with one or more other proteins. Whether the structure defined by the WSXWS motif has any functional role in hematopoietin receptors has not been determined, and there is no evidence to suggest that the WSXWS motif in these other proteins would have a similar function, if it has any at all. An interesting experiment would be to synthesize a WSXWS peptide and test it for its capacity to inhibit ligand binding to hematopoietin receptors, or to block the interactions or function of the above mentioned proteins.

What might be the reason for the conservation of primary amino acid sequence (and presumably tertiary structure) between these receptors? One possibility is that it might reflect the interaction of these receptors with other cell surface molecules that are involved in receptor function. The IL-6 receptor has been shown to be associated with another cell-surface glycoprotein, gp130, that mediates signal transduction[20]. Perhaps this kind of association takes place with all the receptors? Alternatively, the common structural features of the receptors may reflect the structures of the ligands that bind to them. The structures of IL-2 and growth hormone have been determined, by X-ray crystallography, to be helix bundle proteins. It has been proposed that IL-4, erythropoietin and IL-6 also have helix bundle structures, so it is possible that all of the ligands will have this type of structure[15,17,19]. Since most, if not all, of these receptors can be truncated proximal to the transmembrane region to produce a soluble extracellular domain that can still bind ligand, it may soon be possible to determine directly the structures of both ligands and receptors.

Structural differences

Having focused on the sequence similarities found in the members of the hematopoietin receptor superfamily, it is worthwhile to emphasize the differences between them. The IL-6 receptor is unique among these receptors in having an N-terminal immunoglobulin-like domain[10]. In addition, the WSXWS sequence in this receptor is located 51 amino acids upstream of the transmembrane domain, compared with 18–22 amino acids in other members of this receptor family. The most striking difference between the receptors is in the lengths of their cytoplasmic domains (Fig. 1), which range from 54 amino acids (GM-CSFR) to 568 amino acids (human IL-4R). There is not much sequence similarity between the cytoplasmic domains of these receptors, with the exception of the IL-2 receptor and erythropoietin receptor[13]. There are, however, some overall similarities in the amino acid composition of this domain. The IL-2 receptor (IL-2R) and IL-4 receptor (IL-4R) have a relatively large proportion of acidic residues, and the cytoplasmic domains of the IL-2, IL-3, IL-4, IL-7 and erythropoietin receptor are rich in proline and serine residues. Whether these similarities in amino acid composition reflect common signal transduction mechanisms remains to be seen. A serine-rich region in the IL-2R β-subunit cytoplasmic domain has been shown to be important for signal transduction[21]. However, there is no a priori reason to expect that receptors that share extracellular sequence homology should transmit an intracellular signal in the same manner. For example, the PDGF receptor and the IL-1 receptor are both members of the immunoglobulin superfamily, but the former has an intrinsic tyrosine kinase activity in its cytoplasmic domain and the latter does not[1,22]. Some factor-

Table I. Amino acid identities and pairwise ALIGN scores between members of the hematopoietin receptor superfamily

	Amino acid identity (%)								ALIGN scores[a]					
Sequence	huIL7R	muIL7R	huIL4R	muIL4R	huIL2Rβ	huIL6R	muIL3R1	muIl3R2	huGMFR	muEPOR	raPRLR	rbPRLR	huGRHR	rbHRHR
huIL7R[b]	–	33.8	5.6	5.9	3.8	4.8	3.1	6.7	2.4	5.6	6.0	8.2	4.3	3.9
muIL7R	67.2	–	5.6	4.1	4.8	4.6	4.0	5.7	3.1	6.9	5.5	7.1	3.6	4.5
huIL4R	17.2	19.3	–	26.9	5.4	8.2	4.9	8.1	5.7	8.4	3.8	5.7	3.1	3.2
muIL4R	17.0	19.1	50.8	–	4.7	6.1	4.9	7.3	4.8	5.7	4.6	5.5	1.7	1.8
huIL2Rβ	18.7	21.1	18.7	21.2	–	5.4	3.4	8.0	7.3	7.3	4.8	4.3	4.4	2.6
huIL6R	19.5	19.1	23.3	21.7	21.5	–	4.5	5.8	6.6	7.0	8.7	9.3	5.5	6.5
muIL3R1	20.1	19.3	22.7	22.7	18.3	22.0	–	7.6	5.7	4.3	6.2	5.4	5.3	5.0
muIL3R2	20.1	22.0	24.6	24.3	23.5	23.1	21.9	–	8.4	9.9	7.6	6.2	5.2	4.5
huGMFR	17.9	18.2	18.4	18.6	18.5	17.0	18.9	20.3	–	9.7	7.1	6.1	5.9	5.3
muEPOR	18.7	20.9	20.5	17.1	19.9	20.4	15.2	22.0	19.6	–	9.4	9.9	6.2	6.8
raPRLR	19.6	20.6	17.5	21.8	19.0	23.7	19.3	19.5	21.9	21.6	–	47.0	13.3	13.3
rbPRLR	25.5	20.6	21.1	20.8	21.9	25.8	18.3	18.5	20.9	23.7	76.9	–	14.4	14.5
huGRHR	22.4	18.7	20.3	20.0	19.5	20.2	22.4	23.8	15.1	18.0	33.3	35.4	–	51.0
rbGRHR	22.6	20.6	19.8	20.3	18.0	21.2	21.5	20.7	17.9	24.2	34.4	35.4	86.3	–

[a]ALIGN score values were calculated as described in Ref. 19. A score >3 is considered to show a significant evolutionary relationship[1].
[b]Abbreviations: huIL7R, human interleukin-7 receptor; muIL7R, murine interleukin-7 receptor; huIL4R, human interleukin-4 receptor; muIL4R, murine interleukin-4 receptor; huIL2Rβ, human interleukin-2 receptor (β-subunit); huIL6R, human interleukin-6 receptor; muIL3R1, murine interleukin-3 receptor one; muIL3R2, murine interleukin-3 receptor two; huGMFR, human granulocyte–macrophage colony-stimulating factor (GM-CSF) receptor; muEPOR, murine erythropoietin receptor; raPRLR, rat prolactin receptor; rbPRLR, rabbit prolactin receptor; huGRHR, human growth hormone receptor; rbGRHR, rabbit growth hormone receptor.

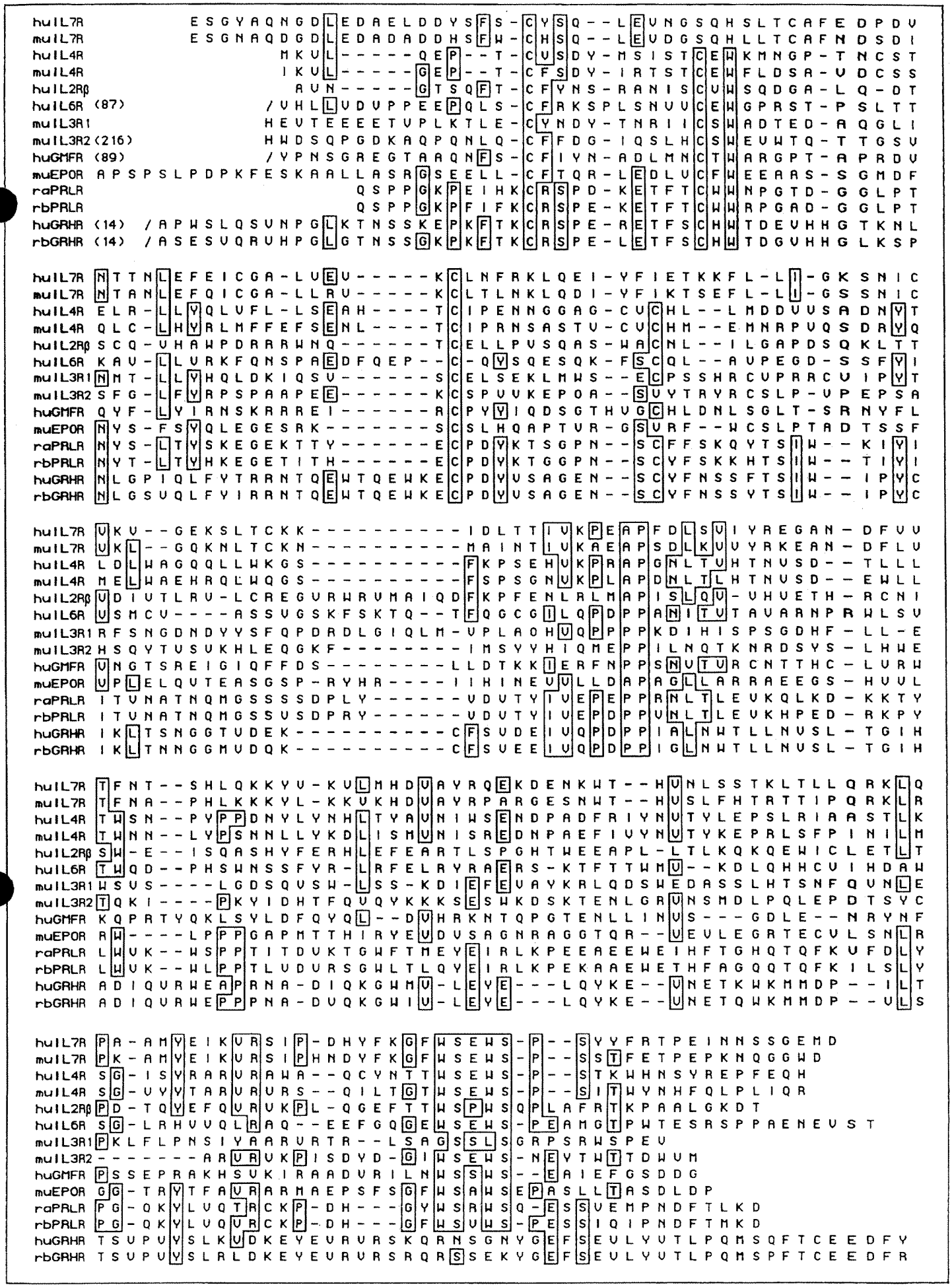
huIL7R
muIL7R
huIL4R
muIL4R
huIL2Rβ
huIL6R (87)
muIL3R1
muIL3R2 (216)
huGMFR (89)
muEPOR
raPRLR
rbPRLR
huGRHR (14)
rbGRHR (14)

Figure 2

Amino acid sequence alignments of the members of the hematopoietin receptor superfamily. Sequences corresponding to the stippled areas in Fig. 1 were aligned as described[19]. Amino acid residues present in six or more of the receptor sequences are boxed. The numbers at the beginning of the sequences indicate the residue within the receptor protein at which the alignment begins. Where no number is shown, the alignment begins at the mature N-terminus of the protein (position 1). For abbreviations, see footnote to Table I.

dependent hematopoietic cell lines are capable of responding to a number of these cytokines by proliferation; these cells can be rapidly adapted from long-term growth in one factor, to growth in another, indicating possible shared signal transduction pathways. There have been reports of increased tyrosine phosphorylation of various proteins in response to addition of these cytokines, including tyrosine phosphorylation of the receptors themselves in the case of IL-2 and IL-3[23,24]. It is also worth noting that IL-3 and GM-CSF can cross compete for binding on some cell lines, but on other cells these cytokines demonstrate entirely independent receptor binding[25]. Whether this reflects an as yet uncloned dual receptor for IL-3 and GM-CSF (the cloned receptors are specific for their respective ligands) or a more complex interaction between the IL-3 and GM-CSF receptors on some cells remains to be determined.

Summary

The discovery of the hematopoietin receptor superfamily has provided a satisfying unity to the field of cytokine research. It has long been speculated that the interleukins and colony stimulating factors might be evolutionarily related, on the basis of the similar sizes of the proteins, similar exon/intron organization of their genes, and common sequence elements involved in control of gene expression. The finding of shared amino acid sequence features in their receptors provides strong support for this view and points the way to understanding cytokine/receptor interactions in molecular detail via mutagenesis and structural analysis. The recent progress in cDNA cloning of receptors for cytokines, neurotransmitters, adhesion molecules and other ligands has shown that certain structural motifs are used by nature over and over again and shuffled around during evolution. The hematopoietin receptors now take their place alongside the other groups of receptors that have been defined by sequence homologies such as the growth factor receptor tyrosine kinases[26], the immunoglobulin superfamily[1], the receptors related to the β-adrenergic receptors[27], and a newly emerging group that includes the nerve growth factor receptor, tumor necrosis factor receptor and CD40[28].

Note added in proof

During the preparation of this manuscript the localization of the disulfide bonds in the extracellular domain of the human growth hormone receptor was published[29]. If the four conserved cysteine residues (indicated as horizontal bars in Fig. 1) are designated as 1–4, starting at the N-terminus of the protein, then Cys1 is paired with Cys2, and Cys3 with Cys4. The next two (non-conserved) cysteine residues are also paired, and the remaining (non-conserved) cysteine located just outside the transmembrane domain (Fig. 2) is unpaired. It will be of great interest to determine if this pattern of disulfide bonding also occurs in the other hematopoietin receptors.

Another recent publication[30] describes the identification of homology between type III fibronectin domains and the C-terminal 90–100 amino acids of the extracellular domains of the hematopoietin receptors. Fibronectin like, type III domains are found in many proteins that mediate adhesion and assembly of complex structures, such as neural adhesion molecules, contactin, fasciclin II and others. It remains to be seen if this domain of the receptors will be involved in ligand binding or in binding to structural components of the cell, either extracellularly or during internalization.

Interestingly, the murine G-CSF receptor has now been cloned[31] and found to belong to the hematopoietin receptor family. The N-terminal portion of the extracellular region consists of an immunoglobulin-like domain (like the IL-6 receptor) followed by a typical hematopoietin receptor domain that mostly closely resembles the prolactin receptor, followed by an additional three fibronectin type III domains before the membrane-spanning region. A similar structure is found for the human G-CSF receptor (A. Larsen *et al.*, unpublished).

Acknowledgements

We thank our colleagues at Immunex for helpful discussions and permission to quote unpublished data. We also thank Linda Troup for preparation of the manuscript.

References

1 Williams, A. F. and Barclay, A. N. (1988) *Annu. Rev. Immunol.* 6, 381–405
2 Mizel, S. B. (1989) *FASEB J.* 3, 2379–2388
3 Nicola, N. A. (1989) *Annu. Rev. Biochem.* 58, 45–77
4 Montgomery, D. W., Zukoski, C. F., Shah, G. N., Buckley, A. R., Pachalczyk, T. and Russell, D. H. (1987) *Biochem. Biophys. Res. Commun.* 145, 692–698
5 Rechler, M. M., Nissley, S. P. and Roth, J. (1987) *N. Engl. J. Med.* 316, 941–943
6 Erslev, A. (1987) *N. Engl. J. Med.* 316, 101–103
7 Leung, D. W., Spencer, S. A., Cachianes, G., Hammonds, R. G., Collins, C., Henzel, W. J., Barnard, R., Waters, M. J. and Wood, W. I. (1987) *Nature* 330, 537–543
8 Boutin, J-M., Jolicoeur, C., Okamura, H., Gagno, J., Edery, M., Shirota, M., Banville, D., Dusanter-Fourt, I., Djiane, J. and Kelley, P. A. (1988) *Cell* 53, 69–77
9 Edery, M., Jolicoeur, C., Levi-Meyrueis, C., Dusanter-Fourt, I., Petridou, B., Boutin, J-M., Lesueur, L., Kelly, P. A. and Djiane, J. (1989) *Proc. Natl Acad. Sci. USA* 86, 2112–2116
10 Yamasaki, K., Taga, T., Hirata, Y., Yawata, H., Kawanishi, Y., Seed, B., Taniguchi, T., Hirano, T. and Kishimoto, T. (1988) *Science* 241, 825–828
11 D'Andrea, A. D., Lodish, H. F. and Wong, G. G. (1989) *Cell* 57, 277–285
12 Hatakeyama, M., Tsudo, M., Minamoto, S., Kono, T., Dori, T., Miyata, T., Miyasaka, M. and Taniguchi, T. (1989) *Science* 244, 551–556
13 D'Andrea, A. D., Fasman, G. D. and Lodish, H. F. (1989) *Cell* 58, 1023–1024
14 Mosley, B., Beckmann, M. P., March, C. J., Idzerda, R. L., Gimpel, S. D., VandenBos, T., Friend, D., Alpert, A., Anderson, D., Jackson, J., Wignall, J. M., Smith, C., Gallis, B., Sims, J. E., Urdal, D. L., Widmer, M. B., Cosman, D. and Park, L. S. (1989) *Cell* 59, 335–348
15 Idzerda, R. L., March, C. J., Mosley, B., Lyman, S. D., VandenBos, T., Gimpel, S. D., Din, W. S., Grabstein, K. H., Widmer, M. B., Park, L. S., Cosman, D. and Beckmann, M. P. (1990) *J. Exp. Med.* 171, 861–873
16 Gearing, D. P., King, J. A. Gough, N. M. and Nicola, N. A. (1989) *EMBO J.* 8, 3667–3676
17 Bazan, J. F. (1989) *Biochem. Biophys. Res. Commun.* 164, 788–795
18 Itoh, N., Yonehara, S., Schreurs, J., Gorman, D. M., Maruyama, K., Ishii, A., Yahara, I., Arai, K-I. and Miyajima, A. (1990) *Science* 247, 324–326
19 Goodwin, R. G., Friend, D., Ziegler, S. F., Jerzy, R., Falk, B. A., Cosman, D., Dower, S. K., March, C. J., Namen, A. E. and Park, L. S. (1990) *Cell* 60, 941–951
20 Taga, T., Hibi, M., Hirata, Y., Yamasaki, K., Yasukawa, K., Matsuda, T., Hirano, T. and Kishimoto, T. (1989) *Cell* 58, 573–581
21 Hatakeyama, M., Mori, H., Dori, T. and Taniguchi, T. (1989) *Cell* 59, 837–845
22 Sims, J. E., March, C. J., Cosman, D., Widmer, M. B., MacDonald, H. R., McMahan, C. J., Grubin, C. E., Wignall, J. M., Jackson, J. L., Call, S. M., Friend, D., Alpert, A. R., Gillis, S., Urdal, D. L. and Dower, S. K. (1988) *Science* 241, 585–589

23 Sharon, M., Gnarra, J. R. and Leonard, W. J. (1989) *J. Immunol.* 143, 2530–2533
24 Isfort, R. J. and Ihle, J. N. (1990) *Growth Factors* 2, 213–220
25 Park, L. S., Friend, D., Price, V., Anderson, D., Singer, J., Prickett, K. S. and Urdal, D. L. (1989) *J. Biol. Chem.* 264, 5420–5427
26 Yarden, Y. and Ullrich, A. (1988) *Annu. Rev. Biochem.* 57, 443–478
27 O'Dowd, B. F., Lefkowitz, R. J. and Caron, M. G. (1989) *Annu. Rev. Neurosci.* 12, 67–83
28 Smith, C. A., Davis, T., Anderson, D., Solam, L., Jerzy, R., Dower, S. K., Cosman, D. and Goodwin, R. G. *Science* (in press)
29 Fuh, G., Mulkerrin, M. G., Bass, S., McFarland, N., Brochler, M., Bourell, J. H., Light, D. R. and Wells, J. A. (1990) *J. Biol. Chem.* 265, 3111–3115
30 Patthy, L. (1990) *Cell* 61, 13–14
31 Fukunaga, R., Ishizaka-Ikeda, E., Seto, Y. and Nagata, S. (1990) *Cell* 61, 341–350

Pheromone response in yeast

In response to peptide pheromones, yeast cells prepare themselves for mating; changes include arrest of the cell cycle and induction of transcription. Proteins involved in this signal transduction pathway include the pheromone receptors, subunits of a G protein, protein kinases and DNA-binding proteins. Understanding of this pathway has been facilitated by yeast genetics, which has allowed the genes encoding all of these proteins to be identified and characterized.

WHEN a CELLS AND α CELLS of the yeast *Saccharomyces cerevisiae* come into close contact, they bind the peptide pheromone released by the opposite cell type. Binding of these pheromones sets in motion a cascade of signaling events that allows the two cells to mate with each other. This fundamental process in a one-celled eukaryote has broad implications for the study in larger eukaryotes of membrane-spanning receptors, G proteins, transcriptional regulators and cell-cycle components.

The mating process in yeast has been studied by investigators approaching the problem and now converging on it from several directions. (1) One major undertaking has been to understand the basis of cell-type specificity: how do cells of opposite mating type arise? (2) A second direction has been the role of the receptors and their downstream components: how does the pheromone signal get transduced? (3) A third approach has focused on the process by which a yeast cell normally completes its cell cycle: how does the mating process disrupt this cycle? The combination of these studies has delineated a pathway of signal response leading from pheromone–receptor interaction to changes in gene transcription and cell-cycle progression. In its overall outline, the pathway has similarities to that used by growth factors in other organisms to regulate cell proliferation.

Control of cell-type specificity

Although **a** and α cells differ in the proteins used for cell–cell recognition (pheromones, receptors and agglutinins), they differ in only a single genetic locus. This mating type locus (*MAT*) encodes transcriptional regulators that control the expression of many unlinked genes. The genes expressed only in the **a** cell (**a**-specific genes) include those for the pheromone **a**-factor and the receptor to the pheromone α-factor; similarly, genes expressed only in the α cell (α-specific genes) include those for α-factor and the receptor to **a**-factor. Another set, the haploid-specific genes, is expressed in both **a** and α cells and encodes a variety of products, including several components of the pheromone response pathway. These gene sets are constitutively transcribed and are responsible for the basic mating apparatus of the haploid cells, which both differentiates the two cell types and allows them to respond to the opposite type. This constitutive (or basal) transcription is dependent on two regulatory networks, as described below: the *MAT* products, and the pheromone response pathway in its unstimulated state. In addition, the transcription of **a**-, α- and many haploid-specific genes is induced by pheromone, and this induction is dependent on the response pathway in its stimulated state.

The *MAT*α allele, contained in α cells, encodes two proteins that regulate α- and **a**-specific gene transcription (the role of these proteins is reviewed in Ref. 1). The *MAT***a** allele, contained in **a** cells, encodes a protein necessary for repression of haploid-specific gene transcription in the **a**/α diploid. Thus the *MAT*-encoded proteins constitute the regulatory switches that determine which sets of cell-type-specific genes are expressed.

STE genes and the signal pathway

The two yeast pheromones, **a**-factor and α-factor, are small peptides which bind to cell surface receptors. α-Factor comprises 13 amino acid residues and is secreted from the cell via the classical secretion pathway. The 12 amino acid **a**-factor is modified at its C-terminus by a methyl ester and a farnesyl group; secretion of **a**-factor is dependent on a protein with homology to the multi-drug resistance P-glycoprotein[2]. Binding of either pheromone leads to several striking changes, including transcriptional induction of a variety of genes required for mating, arrest in the G_1 phase of the cell cycle at a position termed Start, and formation of cell surface projections (reviewed in Ref. 3). Cells displaying these morphological changes have been termed 'shmoos'. The pheromone-treated cells are able to fuse with each other to form a transient heterokaryon, which is followed by fusion of the two haploid nuclei to yield a single diploid nucleus. The resultant zygote is then able to resume vegetative growth (Fig. 1).

Since response to pheromone is a requirement for mating, numerous attempts have been made to identify components involved in the signaling process by identifying mutants unable to mate. Two of the earliest genes identified, *STE2* and *STE3* (*ste* indicates that the mutant is sterile) affect mating in only one cell type; *STE2* is required only by **a** cells and *STE3* only by α cells.

S. Fields is at the Department of Microbiology, State University of New York at Stony Brook, Stony Brook, NY 11794, USA.

A soluble form of CD4 (T4) protein inhibits AIDS virus infection

Keith C. Deen*, J. Steven McDougal†, Richard Inacker*, Gail Folena-Wasserman*, Jim Arthos‡, Jonathan Rosenberg*, Paul Jay Maddon§, Richard Axel‖¶ & Raymond W. Sweet*

* Smith Kline and French Laboratories, King of Prussia, Pennsylvania 19406, USA
† Centers for Disease Control, Atlanta, Georgia 30333, USA
‡ Department of Human Genetics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA
§ Department of Biochemistry and Molecular Biophysics and ‖ Howard Hughes Medical Institute, Columbia University, 722 West 168th St, New York, New York 10032, USA

CD4 (T4) is a glycoprotein of relative molecular mass 55,000 (M_r 55K) on the surface of T lymphocytes which is thought to interact with class II MHC (major histocompatibility complex) molecules, mediating efficient association of helper T cells with antigen-bearing targets[1–3]. The CD4 protein is also the receptor for HIV, a T-lymphotropic RNA virus responsible for the human acquired immune deficiency syndrome (AIDS) (refs 4–7). To define the mechanisms of interaction of CD4 with the surface of antigen-presenting cells and with HIV, we have isolated the CD4 gene and expressed this gene in several different cellular environments[7,8]. Here we describe an efficient expression system in which a recombinant, soluble form of CD4 (sCD4) is secreted into tissue culture supernatants. This sCD4 retains the structural and biological properties of CD4 on the cell surface, binds to the envelope glycoprotein (gp110) of HIV and inhibits the binding of virus to CD4⁺ lymphocytes, resulting in a striking inhibition of virus infectivity.

The CD4 molecule is comprised of an N-terminal hydrophobic signal sequence, four extracellular domains with homology to immunoglobulin variable and joining regions, a hydrophobic transmembrane domain, and a highly charged cytoplasmic segment[8,9]. To obtain large quantities of the extracellular segment of CD4, we introduced a termination codon into the CD4 complementary DNA[8] at the extracellular boundary of the transmembrane domain, expecting that the truncated protein would be secreted in a soluble form. This sCD4 cDNA was inserted between the SV40 early promoter and the bovine growth hormone (BGH) polyadenylation site (Fig. 1*a*). The sCD4 expression cassette was linked to a mouse dihydrofolate reductase (*dhfr*) expression cassette[10] to permit co-amplification of the sCD4 gene on exposure to increasing concentrations of the *dhfr* inhibitor methotrexate (MTX). This vector, psT4DHFR (Fig. 1*a*), was transfected into the *dhfr*-deficient CHO cell line, DXB11 (ref. 11), and transformants were selected and exposed to progressively increasing concentrations of methotrexate[12]. Supernatants from clones were monitored for the expression of sCD4 by one of three independent assays: (1) immunoprecipitation from ^{35}S-labelled cultures with CD4-specific monoclonal antibodies; (2) Western blot analysis using a rabbit anti-CD4 polyclonal antibody developed against a denatured CD4 protein produced in bacteria; and (3) competition enzyme-linked immunosorbent assay (ELISA) performed with immobilized sCD4.

^{35}S-methionine- and ^{35}S-cysteine-labelled sCD4 from culture supernatants was specifically immunoprecipitated by each of six anti-CD4 monoclonal antibodies but not by control antibodies (Fig. 1*b*). In contrast, when culture supernatants were analysed under denaturing conditions by Western blot analysis, sCD4 was not recognized by any of the six anti-CD4 monoclonal

¶ To whom correspondence should be addressed.

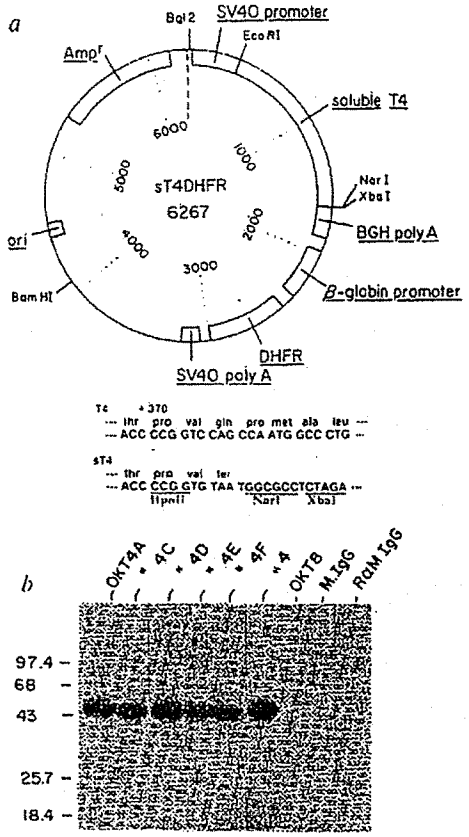

Fig. 1 *a*, Plasmid psT4DHFR is a pUC18 derivative containing base pairs (bp) 1–1,257 of the CD4 cDNA clone pT4B[8] which encodes the leader and extracellular segment of CD4. This sCD4 cDNA is inserted between an SV40 early promoter and a synthetic linker containing a TAA termination codon (inset) followed by the polyadenylation region of the bovine growth hormone gene. The sCD4 expression cassette is linked to a mouse *dhfr* expression cassette consisting of the β-globin promoter, mouse *dhfr* coding sequence and the SV40 polyadenylation region. *b*, Epitope characterization of sCD4. Cell lines were metabolically labelled with ^{35}S-methionine and cysteine and the supernatants were immunoprecipitated with monoclonal antibodies to CD4 (OKT4, T4A, T4C-F), CD8, and control antibodies (mouse IgG (M. IgG) and rabbit anti-mouse IgG (RαM1gG)). Regular molecular masses (M_r) are given in thousands.

Methods. *a*, Plasmid psT4DHFR was constructed in pUC18 and contains the following sequence elements isolated from previously reported plasmids: SV40 early promoter[10], SV40 poly(A) early region[10], bovine growth hormone poly(A) region[10], mouse β-globin promoter[14], mouse *dhfr* coding region[15] and bp 1–1257 of the CD4 cDNA clone pT4B[8]. The full-length CD4 cDNA was truncated at a *Hpa*II site (bp 1,252) and ligated to a synthetic linker (inset) which restored bp 1,253–1,257 of the CD4 coding sequence and placed a TAA termination codon after bp 1,257. DXB-11 cells, a *dhfr*⁻ CHO cell line[11], were transfected by calcium phosphate precipitation[16] with 10 or 30 μg of psT4DHFR one day after seeding. Transformants were selected in nucleoside-free medium and the transforming sequences were amplified through rounds of increasing MTX concentrations[12]. *b*, Cultures containing 1×10^6 cells per 60 mm culture dish were labelled for 16 h at 37 °C in 1.5 ml methionine and cysteine free F12 medium containing ITS (Collaborative), glutamine, 170 μCi ml^{-1} ^{35}S-methionine and 30 μCi ml^{-1} ^{35}S-cysteine (ICN). Clarified media was precleared, incubated with 5 μg OKT4, T4A, T4C-T4F, OKT8 (P. Rao, Ortho), mouse IgG (Cooper), or rabbit anti-mouse IgG (Cooper) for 30 min at 4 °C, then incubated with protein A-Sepharose beads. The beads were washed twice, boiled for 5 min in 20 μl sample buffer and analysed by electrophoresis through a 12.5% SDS-polyacrylamide gel under reducing conditions.

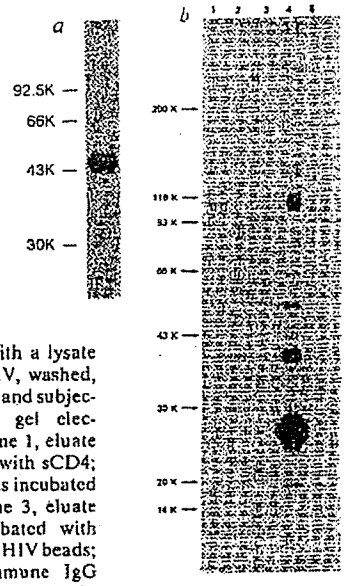

Fig. 2 *a*, SDS-PAGE profile of purified sCD4. Relative molecular masses (M_r) are given in kilodaltons. *b*, Coprecipitation of HIV gp110 with sCD4. Supernatant containing sCD4 and control supernatant from untransformed DXB-11 cells were adsorbed to Sepharose beads coated with OKT4, control antibody MOPC141 (isotype matched to OKT4), or human anti-HIV IgG. The beads were washed, mixed with a lysate of ^{35}S-methionine-labelled HIV, washed, and bound material was eluted and subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Lane 1, eluate from OKT4 beads incubated with sCD4; lane 2, eluate from OKT4 beads incubated with control supernatant; lane 3, eluate from MOPC141 beads incubated with sCD4; lane 4, eluate from anti-HIV beads; lane 5, eluate from non-immune IgG beads. Relative molecular masses are given.

Methods. *a*, Purified sCD4 (3 μg) was electrophoresed through a 12.5% polyacrylamide gel and visualized by silver stain. *b*, A lysate of extracellular HIV was internally labelled with ^{35}S-methionine and ^{35}S-cysteine as previously described[6]. Sepharose beads coupled to OKT4 monoclonal antibody, MOPC141 paraprotein (isotype matched to OKT4), human anti-HIV IgG and non-immune human IgG were prepared as previously described[6]. OKT4 beads and MOPC141 beads (10 μl) were incubated for 2 h at 4 °C with ~5 μg of sCD4, or control supernatant from untransformed DXB-11 cells, then washed once with lysing buffer. The ^{35}S-labelled HIV (preadsorbed) was added and the beads were incubated for 3 h at 4 °C. Washing, elution of bound material and SDS-PAGE under reducing conditions were performed as previously described[6].

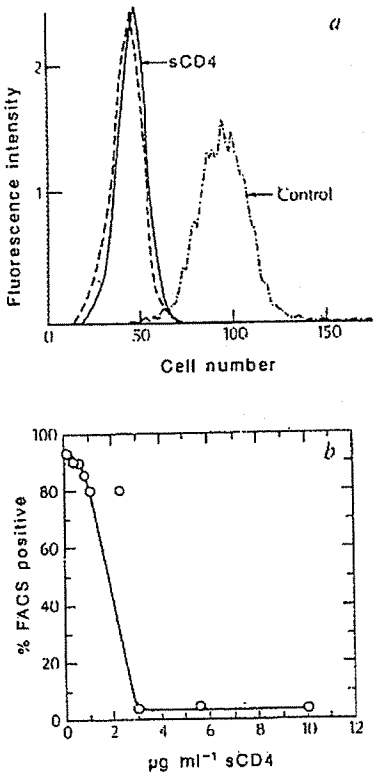

Fig. 3 *a*, sCD4 inhibits HIV binding to CD4$^+$ CEM T cells. Cells were incubated with buffer (- - - -), HIV precincubated with sCD4 (——), or with HIV preincubated with concentrate control supernatant from untransformed DXB-11 (—·—·), washed, exposed to fluorescein-conjugated anti-HIV antibody, and analysed by cytofluorometry as previously described[11]. A fluorescence histogram (cell number against fluorescence intensity) is shown. *b*, Plot of percentage of positive cells by flow cytometry versus concentration of sCD4.

antibodies (which recognize only native protein), but was readily detected by rabbit anti-CD4 polyclonal serum raised against a denatured form of CD4 produced in bacteria (not shown). Several clones synthesize >3 pg of sCD4 per cell in 24 hours, giving ~40 mg of sCD4 per litre in high-density suspension cultures over 4 days. We have purified large quantities of sCD4 to >95% purity (Fig. 2*a*), and this purified protein shows the same pattern of antibody reactivity as described for sCD4 in culture supernates. These data suggest that sCD4 maintains a configuration analogous to the native extracellular domain of CD4 on the lymphocyte surface.

We next examined whether sCD4 associates with HIV envelope glycoprotein gp110. Purified sCD4 (~5 μg) was adsorbed to Sepharose beads coated with OKT4 or control antibody. OKT4 associates with the CD4 molecule without inhibiting the ability of CD4 to interact with virus[4-7]. The beads were mixed with a lysate of ^{35}S-methionine-labelled HIV. Only the 110K envelope glycoprotein is coprecipitated by the complex of sCD4 with OKT4 (Fig. 2*b*, lane 1); no viral proteins are precipitated by OKT4 beads without sCD4 or in the presence of control supernatants from the untransfected CHO cells (Fig. 2*b*, lane 2). Furthermore, no viral protein is precipitated if Sepharose beads coated with control mouse immunoglobulin (isotype matched to OKT4) are incubated with sCD4 (Fig. 2*b*, lane 3). Thus the soluble CD4 protein, without other T-cell surface components can specifically associate with the envelope glycoprotein of the AIDS virus.

To examine whether sCD4 abolishes HIV binding to the surface of CD4$^+$ cells, CEM T cells (which carry CD4) were exposed to HIV in the presence or absence of sCD4. After viral adsorption, the cells were washed, exposed to fluorescein-conjugated anti-HIV antibody, and analysed by flow cytometry (Fig. 3*a*)[6,7]. HIV binds efficiently to CEM cells; but if the virus is preincubated with sCD4 the binding is abolished; 3 ng of purified sCD4 is sufficient to inhibit the binding of 100 ng of viral protein (Fig. 3*b*). We estimate that if the envelope glycoprotein comprises 5% of the total viral protein, then a 2:1 molar ratio of CD4 to gp120 can completely inhibit HIV binding to CD4$^+$ cells.

Finally, we have examined the ability of soluble CD4 to inhibit the infection of CD4$^+$ cells by HIV. Phytohaemagglutinin (PHA)-stimulated human lymphocytes were exposed to serial ten-fold dilutions of an HIV inoculum in the presence or absence of sCD4, washed and plated in microculture. The frequency of infected culture was determined using an immunoassay 4, 8 and 12 days after exposure to virus[13] (Fig. 4*a*). We define the infectious virus titre, ID-50, as the reciprocal of the dilution required to infect 50% of the exposed cell cultures at day 12 (Fig. 4*b*). In the absence of sCD4, the ID-50 observed with our viral inoculum is ~10^5. However, in the presence of 8 μg ml^{-1} purified sCD4, the infectivity is diminished to an ID-50 of $10^{1.5}$ (Fig. 4*b*). This dramatic reduction in infectivity is observed throughout the course of infection. As a control for nonspecific inhibition or toxic effects, sCD4 was added to cultures 18 h after the initial exposure to virus. These cultures show only a 1 log inhibition in the ID-50 (Fig. 4*b*), probably resulting from inhibition of virus spread after the initial inoculation. Thus, the inhibition of virus infection by sCD4 is likely to be a result of

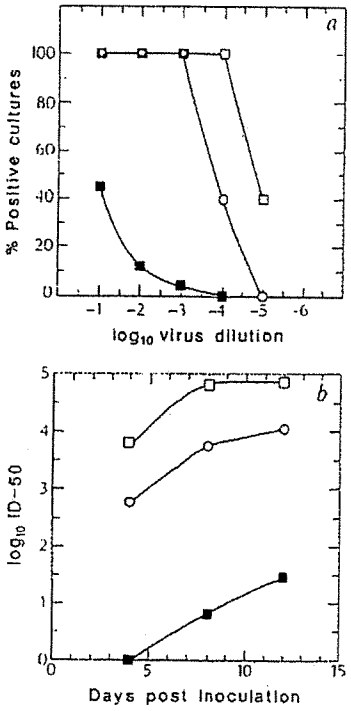

Fig. 4 Inhibition of HIV infectivity by sCD4. *a*, Plot of percentage of cultures positive for HIV at day 8 versus dilution of virus inoculum; *b*, Plot of ID-50 (reciprocal of virus dilution at which 50% of cultures are positive) at days 4, 8 and 12. Infectivity titration of an HIV inoculum (ID-50 assay) was performed as previously described[17]. Briefly, serial 10-fold dilutions of virus inoculum are incubated with indicator cells (PHA-stimulated human lymphocytes) for 18 h. The cells are then washed and plated in microculture (1×10^5 cells per culture, 10 cultures per dilution). At days 4, 8 and 12, supernatants are tested for HIV by the antigen capture assay[13]. ID-50 titrations were performed in media containing 8.6 μg ml^{-1} sCD4 which was added to the HIV dilution 30 min before inoculation of cells and maintained in the culture media throughout the experiment (■—■), or in media containing sCD4 introduced after the initial 18 h inoculum (○—○) (delayed addition control), or in media without sCD4 (□—□), control).

the specific association of sCD4 with gp110, inhibiting the interaction of virus with CD4$^+$ cells. We estimate that our viral preparations of 10^4 infectious particles per ml actually contain 10^8 particles per ml. If each particle contains 1,000 envelope glycoproteins, the 3.5 log inhibition we observe is obtained at a ratio of 1,000 CD4 molecules per molecule of envelope glycoprotein.

The ability of sCD4 to bind gp110 and inhibit viral infection *in vitro* immediately suggests the potential use of sCD4 as an antiviral agent in AIDS patients. Although significant variance exists among the different HIV isolates, each appears to use CD4 as a cellular receptor. Thus, sCD4 is likely to be a universal inhibitor of viral infection. Although sCD4 is extremely effective *in vitro*, its use in man will depend on the pharmacokinetics of sCD4, its immunogenicity, its effects on the cellular immune response and the clinical significance of inhibiting viral spread in infected individuals. Whatever the efficacy of sCD4 as a therapeutic in AIDS, this reagent should allow a more precise understanding of the structural basis for the association of CD4 with the HIV envelope glycoprotein and with the surface of antigen-presenting cells.

We thank S. P. Cort, M. S. Kennedy, A. Kandil, C. Silverman and A. Trunch for assistance during the project, M. Rosenberg for support and encouragement, Ortho Pharmaceuticals for providing OKT4 and Brian Soda for preparation of the figures. This work was supported by the Howard Hughes Medical Institute (R.A.) and by a grant from the National Institutes of Health (P.J.M. and R.A.).

Received 28th October; accepted 26 November 1987.

1. Gay, D. *et al. Nature* 328, 626–629 (1987).
2. Sleckman, B. D. *et al. Nature* 328, 351–353 (1987).
3. Thomas, Y., Rognozinski, L. & Chess, L. *Immun. Rev.* 74, 113–128 (1985).
4. Dalgleish, A. G. *et al. Nature* 312, 763–766 (1984).
5. Klatzman, D. *et al. Nature* 312, 767–768 (1984).
6. McDougal, J. S. *et al. Science* 231, 382–385 (1986).
7. Maddon, P. J. *et al. Cell* 47, 333–348 (1986).
8. Maddon, P. J. *et al. Cell* 42, 93–104 (1985).
9. Maddon, P. J. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
10. Pfarr, D. S., Sathe, G. & Reff, M. F. *DNA* 4, 461–467 (1985).
11. Urlaub, G., Kas, E., Carothers, A. M. & Chasin, L. A. *Cell* 33, 405–412 (1983).
12. Kaufman, R. J., Sharp, P. A. & Latt, S. A. *Molec. cell. Biol.* 3, 699–711 (1983).
13. McDougal, J.-S. *et al. J. immun. Meth.* 76, 171–183 (1985).
14. Berg, P. E. *et al. Molec. cell. Biol.* 3, 1246–1254 (1983).
15. Subramani, S., Mulligan, R. & Berg, P. *Molec. cell. Biol.* 1, 854–864 (1981).
16. Wigler, M. *et al. Cell* 11, 223–232 (1977).
17. McDougal, J. S. *et al. J. Immun.* 135, 3151–3162 (1985).

Soluble CD4 molecules neutralize human immunodeficiency virus type 1

André Traunecker, Wolfgang Lüke* & Klaus Karjalainen

Basel Institute for Immunology, Grenzacherstrasse 487, CH-4058 Basel, Switzerland
* Deutsches Primatenzentrum, Abteilung Virologie und Immunologie, Kellnerweg 4, Göttingen 3400, FRG

Human immunodeficiency virus (HIV) infection can bring about total collapse of the immune system[1,2] by infecting helper T lymphocytes which express CD4, the molecule which mediates interaction between the cell surface and viral envelope glycoprotein gp120 (refs 3–10). HIV apparently escapes the effects of neutralizing antibodies *in vivo* by generating new variants which must still interact with CD4 to maintain a cycle of infection[11–14]. One route to block HIV infection, therefore, could use solubilized CD4 protein to inhibit attachment of the virus to its target cell. We have used recombinant DNA techniques to generate soluble forms of CD4, and show here that these are potent inhibitors of HIV infection *in vitro*.

Two different chimaeric genes based on immunoglobulin-expression systems were made (Fig. 1). The construct HT4-Y1 codes for the N-terminal immunoglobulin-like region and three following domains of CD4 which comprise the whole extracellular portion; the other, HT4-X6, codes for the immunoglobulin-like and second domains only[15,16]. In both cases the C-terminal part of the protein consists of the immunoglobulin κ light-chain constant region.

The introduction of these constructs into myeloma cells (Fig. 2) resulted in the secretion of the two chimaeric proteins of predicted apparent molecular weight: HT4-Y1 and HT4-X6 transformants produced molecules of relative molecular mass (M_r) 60,000 (60K) and 30K respectively, as detected by Western blotting with anti-κ antibodies (Fig. 2*a*). Furthermore, immunoprecipitations of the supernatants with two different monoclonal anti-CD4 antibodies (Fig. 2*b*) showed that these secreted proteins retained at least some of their original conformation.

In addition, the soluble CD4 proteins could bind at low concentration to purified HIV-envelope protein gp120 in solid-phase assays, consistent with a high-affinity interaction[10] (Fig. 3). The lower intensity of the signal generated by the 'complete' CD4 (Y124) compared to the truncated one (X6-10) is probably due to a concentration difference in the supernatants (20 and